Automotive Math and Motor Control Library Set for NXP S32K14x devices

Accuracy of Floating-Point Functions

Document Number: S32K14XMCFLTACC
Rev. 3

Contents

Section number **Title** **Page**

Chapter 1
Revision History

Chapter 2
Introduction

Chapter 3
Library Functions

Section number	Title	Page

Section number	Title	Page

Section number	Title	Page

Section number	Title	Page

Section number	Title	Page

Section number	Title	Page

Section number	Title	Page

Chapter 1
Revision History

Table 1-1. Revision History

Revision	Date	Author	Description
1.0	30/09/2015	Petr Zelinka	Initial version.
2.0	1/04/2016	Petr Zelinka	Updated GFLIB_AtanYX_FLT and GFLIB_AtanYXShifted_FLT.
3.0	31/12/2016	Petr Zelinka	Added AMCLIB_BemfObsrvDQ_FLT and AMCLIB_TrackObsrv_FLT functions.

Chapter 2
Introduction

Applications utilizing floating-point calculations tend to exhibit hard-to-predict behavior due to intricacies of finite-precision implementation of numerical algorithms. Such uncertainty cannot be ignored in critical systems. The risk of accidents resulting from numerical errors can be alleviated if the limitations of the algorithms are well understood.

Standard C libraries are typically designed to provide correctly rounded results for all functions (i.e. error less than ±0.5 ulp). Maintaining such accuracy is prohibitively slow for embedded applications. The Automotive Math and Motor Control Library Set for NXP S32K14x devices is tailored for embedded automotive systems requiring maximum processing speed and useful accuracy of results.

The aim of this document is to specify the guaranteed accuracy of floating-point functions contained in the Automotive Math and Motor Control Library Set for NXP S32K14x devices. The accuracy generally depends on specific numerical cases and hence is specified for all combinations of floating-point inputs, including all parameters and state variables accessed via input pointers. The guaranteed accuracy is expressed in terms of worst-case output error bounds for each specific case.

The accuracy criteria described in this document are also available in the form of Matlab® scripts suitable for automated verification systems.

2.1 About this Manual

This document employs the following typographical conventions:

Table 2-1. Typographical Conventions

Typographical Style	Example
Symbols in capital italic font represent sets.	*M*
Symbols in bold italic font represent vectors.	***fltTable***
symbols in italic font represent scalars.	*fltIn*
Subscript is used for indexing of vector elements.	$fltTable_n$

2.2 Acronyms and Definitions

Table 2-2. Acronyms and Definitions

Term	Definition
API	Application Programming Interface.
Inf	Floating-point special value "infinity".
NaN	Floating-point special value "not a number".
nmax	Maximum negative normalized single precision floating-point value.
pmax	Maximum positive normalized single precision floating-point value.
ulp	Unit in the last place. Used as a unit of measurement of floating-point calculation error.

2.3 Reference List

Table 2-3. Reference List

#	Title
1	IEEE 754-2008 IEEE Standard for Floating-Point Arithmetic, 2008
2	Power ISA , Version 2.06 Revision B, July 23, 2010
3	ARMv7-M Architecture Reference Manual, Issue Derrata 2010_Q3, 2010
4	ISO/IEC 9899:1999 Programming Language – C, 1999
5	Automotive Math and Motor Control Library Set for NXP S32K14x devices User's Guide, revision 5

2.4 Common Definitions

This document uses the following set builder notation for input values of library functions:

$\{(fltIn_1, fltIn_2) \in X \mid X \cap M = \varnothing \wedge X \cap D \neq \varnothing\}$ denotes a set X of one or more combinations of input values $fltIn_1$, $fltIn_2$, such that X does not contain elements of M and contains at least one element of D. $X \subset Z$ where Z is a set of all possible combinations of all floating-point values which can appear on the inputs.

The guaranteed accuracy of results is defined for each subset of input values. Whenever two subsets overlap, the one with higher allowed error takes precedence.

Unless otherwise stated, variables are to be taken as real numbers. Whenever a variable is compared to Inf, -Inf, or NaN, a single precision floating-point representation of such variable in round-to-nearest mode is considered.

Unless otherwise stated, inputs to library functions are considered exactly representable in single precision floating-point format.

Let $N = \langle -(2 - 2^{-23}) \cdot 2^{127}, -2^{-126} \rangle \cup \langle 2^{-126}, (2 - 2^{-23}) \cdot 2^{127} \rangle$, i.e. a set of real numbers which are represented by normalized numbers in single precision floating-point format.

Let $D = \langle -2^{-126}, -2^{-149} \rangle \cup \langle 2^{-149}, -2^{-126} \rangle$, i.e. a set of real numbers which are represented by denormalized numbers in single precision floating-point format, excluding zero.

Let M = {NaN, Inf, -Inf}, i.e. a set of non-numerical floating-point values.

Let $pmax = (2 - 2^{-23}) \cdot 2^{127}$, i.e. the maximum positive normalized single precision floating-point value.

Let $nmax = -(2 - 2^{-23}) \cdot 2^{127}$, i.e. the maximum negative normalized single precision floating-point value.

Function ceil(x) provides x rounded towards positive infinity.

Function fe(x) provides the unbiased exponent of a single precision floating-point representation of x.

Function fix(x) provides x rounded towards zero.

Function floor(x) provides x rounded towards negative infinity.

Function max(a,b,c,...) provides the maximum value of {a, b, c, …}.

Function min(a,b,c,...) provides the minimum value of {a, b, c, …}.

Function sign(x) provides 1 if $x > 0$, 0 if $x = 0$, and -1 if $x < 0$.

If x lies between two finite consecutive single precision floating-point numbers a and b without being equal to one of them, then ULP(x) = $|b - a|$, otherwise ULP(x) is the distance between two finite single precision floating-point numbers nearest x. ULP(NaN) = NaN.

The error of floating-point results is measured in units of ulp by comparing the actual *result* with theoretical exact *refResult* as follows:

$$e = \frac{refResult - result}{\mathrm{ULP}(refResult)}\,[\mathrm{ulp}]$$

Equation 1

E.g. e = 10 ulp describes a result that deviates from exact result by ten times the absolute value of the least-significant mantissa bit of a single precision floating-point representation of the exact result.

All error measurements assume IEEE 754-2008 binary floating-point arithmetic with round-to-nearest rounding mode and default results mode in case of Power ISA implementation.

Chapter 3
Library Functions

This sections provides error bounds for individual functions in the Automotive Math and Motor Control Library Set for NXP S32K14x devices.

3.1 Function AMCLIB_BemfObsrvDQ_FLT

3.1.1 Declaration

```
tFloat AMCLIB_BemfObsrvDQ_FLT(const SWLIBS_2Syst_FLT *const pIAB, const SWLIBS_2Syst_FLT
*const pUAB, tFloat fltVelocity, tFloat fltPhase, AMCLIB_BEMF_OBSRV_DQ_T_FLT *const pCtrl);
```

3.1.2 Arguments

Table 3-1. AMCLIB_BemfObsrvDQ_FLT arguments

Type	Name	Direction	Description
const SWLIBS_2Syst_FLT *const	pIAB	**input**	Pointer to the structure with Alpha/Beta current components [A].
const SWLIBS_2Syst_FLT *const	pUAB	**input**	Pointer to the structure with Alpha/Beta voltage components [V].
tFloat	fltVelocity	**input**	Estimated electrical angular velocity [rad/s].
tFloat	fltPhase	**input**	Estimated rotor flux angle [rad], must be in range [-π, π].
AMCLIB_BEMF_OBSRV_DQ_T_FLT *const	pCtrl	**input, output**	Pointer to the structure with BEMF observer coefficients.

3.1.3 Worst-Case Error Bounds

Let (*pIABArg1*, *pIABArg2*, *pUABArg1*, *pUABArg2*, *fltVelocity*, *pPhaseErr*, *pIObsrvArg1*, *pIObsrvArg2*, *DfltCC1sc*, *DfltCC2sc*, *DfltAcc*, *DfltInErrK1*, *DfltUpperLimit*, *DfltLowerLimit*, *DfltCC1sc*, *DfltCC2sc*, *DfltAcc*, *DfltInErrK1*, *DfltUpperLimit*, *DfltLowerLimit*, *pIObsrvIn_1Arg1*, *pIObsrvIn_1Arg2*, *fltIGain*, *fltUGain*, *fltWIGain*, *fltEGain*) ∈ *X* be a vector of inputs to AMCLIB_BemfObsrvDQ_FLT,

pIObsrvArg1, *pIObsrvArg2*, *DfltAcc*, *QfltAcc*, *DfltInErrK1*, *QfltInErrK1*, *pIObsrvIn_1Arg1*, *pIObsrvIn_1Arg2* inputs represent real numbers with an error of max. ±0.5 ulp,

$ir_1 = pUABArg1 \cdot \cos(pPhaseErr)$,

$me_1 = 668$,

$ir_2 = pUABArg2 \cdot \sin(pPhaseErr)$,

$me_2 = 668$,

$ie_1 = \max(\mathrm{fe}(|ir_1| + me_1 \cdot \mathrm{ULP}(ir_1)), \mathrm{fe}(|ir_2| + me_2 \cdot \mathrm{ULP}(ir_2)))$,

$ir_3 = ir_1 + ir_2$,

$ie_2 = \mathrm{fe}(ir_3)$,

$$cb_1 = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me_3 = 1 + 1336 \cdot 2^{cb_1}$,

$ir_4 = pUABArg2 \cdot \cos(pPhaseErr)$,

$me_4 = 668$,

$ir_5 = -pUABArg1 \cdot \sin(pPhaseErr)$,

$me_5 = 668$,

$ie_3 = \max(\mathrm{fe}(|ir_4| + 668 \cdot \mathrm{ULP}(ir_4)), \mathrm{fe}(|ir_5| + 668 \cdot \mathrm{ULP}(ir_5)))$,

$ir_6 = ir_4 + ir_5$,

$ie_4 = \mathrm{fe}(ir_6)$,

$$cb_2 = \begin{cases} 0, & ie_3 - ie_4 \le 0 \\ ie_3 - ie_4, & ie_3 - ie_4 > 0 \end{cases},$$

$$me_6 = 1 + 1336 \cdot 2^{cb_2},$$

$$ir_7 = pIABArg1 \cdot \cos(pPhaseErr),$$

$$me_7 = 668,$$

$$ir_8 = pIABArg2 \cdot \sin(pPhaseErr),$$

$$me_8 = 668,$$

$$ie_5 = \max(\mathrm{fe}(|ir_7| + 668 \cdot \mathrm{ULP}(ir_7)), \mathrm{fe}(|ir_8| + 668 \cdot \mathrm{ULP}(ir_8))),$$

$$ir_9 = ir_7 + ir_8,$$

$$ie_6 = \mathrm{fe}(ir_9),$$

$$cb_3 = \begin{cases} 0, & ie_5 - ie_6 \le 0 \\ ie_5 - ie_6, & ie_5 - ie_6 > 0 \end{cases},$$

$$me_9 = 1 + 1336 \cdot 2^{cb_3},$$

$$ir_{10} = pIABArg2 \cdot \cos(pPhaseErr),$$

$$me_{10} = 668,$$

$$ir_{11} = -pIABArg1 \cdot \sin(pPhaseErr),$$

$$me_{10} = 668,$$

$$ie_7 = \max(\mathrm{fe}(|ir_{10}| + 668 \cdot \mathrm{ULP}(ir_{10})), \mathrm{fe}(|ir_{11}| + 668 \cdot \mathrm{ULP}(ir_{11}))),$$

$$ir_{12} = ir_{10} + ir_{11},$$

$$ie_8 = \mathrm{fe}(ir_{12}),$$

$$cb_4 = \begin{cases} 0, & ie_7 - ie_8 \le 0 \\ ie_7 - ie_8, & ie_7 - ie_8 > 0 \end{cases},$$

$$me_{12} = 1 + 1336 \cdot 2^{cb_4},$$

$$ie_9 = \max(\mathrm{fe}(|ir_9| + me_9 \cdot \mathrm{ULP}(ir_9)), \mathrm{fe}(|pIObsrvArg1| + 0.5 \cdot \mathrm{ULP}(pIObsrvArg1))),$$

$$ir_{13} = ir_9 - pIObsrvArg1,$$

$$ie_{10} = \mathrm{fe}(ir_{13}),$$

$$cb_5 = \begin{cases} 0, & ie_9 - ie_{10} \le 0 \\ ie_9 - ie_{10}, & ie_9 - ie_{10} > 0 \end{cases},$$

$$me_{13} = 0.5 + (me_9 + 0.5) \cdot 2^{cb_5},$$

$$ie_{11} = \max(\mathrm{fe}(|ir_{12}| + me_{12} \cdot \mathrm{ULP}(ir_{12})), \mathrm{fe}(|pIObsrvArg2| + 0.5 \cdot \mathrm{ULP}(pIObsrvArg2))),$$

$$ir_{14} = ir_{12} - pIObsrvArg2,$$

$$ie_{12} = \mathrm{fe}(ir_{14}),$$

$$cb_6 = \begin{cases} 0, & ie_{11} - ie_{12} \le 0 \\ ie_{11} - ie_{12}, & ie_{11} - ie_{12} > 0 \end{cases},$$

$$me_{14} = 0.5 + (me_{12} + 0.5) \cdot 2^{cb_6},$$

$$ir_{15} = ir_{13} \cdot DfltCC1sc,$$

$$me_{15} = 2 \cdot me_{13},$$

$$ir_{16} = DfltInErrK1 \cdot DfltCC2sc,$$

$$me_{16} = 1,$$

$$ie_{13} = \max(\mathrm{fe}(|ir_{15}| + me_{15} \cdot \mathrm{ULP}(ir_{15})), \mathrm{fe}(|ir_{16}| + \mathrm{ULP}(ir_{16}))),$$

$$ir_{17} = ir_{15} + ir_{16},$$

$$ie_{14} = \mathrm{fe}(ir_{17}),$$

$$cb_7 = \begin{cases} 0, & ie_{13} - ie_{14} \le 0 \\ ie_{13} - ie_{14}, & ie_{13} - ie_{14} > 0 \end{cases},$$

$$me_{17} = 1 + (me_{15} + 1) \cdot 2^{cb_7},$$

$$ir_{18} = ir_{14} \cdot QfltCC1sc,$$

$$me_{18} = 2 \cdot me_{14},$$

$$ir_{19} = QfltInErrK1 \cdot QfltCC2sc,$$

$$me_{19} = 1,$$

$$ie_{15} = \max(\mathrm{fe}(|ir_{18}| + me_{18} \cdot \mathrm{ULP}(ir_{18})), \mathrm{fe}(|ir_{19}| + \mathrm{ULP}(ir_{19}))),$$

$$ir_{20} = ir_{18} + ir_{19},$$

$$ie_{16} = \mathrm{fe}(ir_{20}),$$

$$cb_8 = \begin{cases} 0, & ie_{15} - ie_{16} \le 0 \\ ie_{15} - ie_{16}, & ie_{15} - ie_{16} > 0 \end{cases},$$

$$me_{20} = 1 + (me_{18} + 1) \cdot 2^{cb_8},$$

$$ie_{17} = \max(\mathrm{fe}(|ir_{17}| + me_{17} \cdot \mathrm{ULP}(ir_{17})), \mathrm{fe}(|DfltAcc| + 0.5 \cdot \mathrm{ULP}(DfltAcc))),$$

$$ir_{21} = ir_{17} + DfltAcc,$$

$$ie_{18} = \mathrm{fe}(ir_{21}),$$

$$cb_9 = \begin{cases} 0, & ie_{17} - ie_{18} \le 0 \\ ie_{17} - ie_{18}, & ie_{17} - ie_{18} > 0 \end{cases},$$

$$me_{21} = \begin{cases} 0, & ir_{21} + \left(0.5 + (me_{17} + 0.5) \cdot 2^{cb_9}\right) \cdot \mathrm{ULP}(ir_{21}) < DfltLowerLimit \\ 0, & ir_{21} - \left(0.5 + (me_{17} + 0.5) \cdot 2^{cb_9}\right) \cdot \mathrm{ULP}(ir_{21}) > DfltUpperLimit \\ 0.5 + (me_{17} + 0.5) \cdot 2^{cb_9}, & \text{otherwise} \end{cases},$$

$$ie_{19} = \max(\mathrm{fe}(|ir_{20}| + me_{20} \cdot \mathrm{ULP}(ir_{20})), \mathrm{fe}(|QfltAcc| + 0.5 \cdot \mathrm{ULP}(QfltAcc))),$$

$$ir_{22} = ir_{19} + QfltAcc,$$

$$ie_{20} = \mathrm{fe}(ir_{22}),$$

$$cb_{10} = \begin{cases} 0, & ie_{19} - ie_{20} \le 0 \\ ie_{19} - ie_{20}, & ie_{19} - ie_{20} > 0 \end{cases},$$

$$me_{22} = \begin{cases} 0, & ir_{22} + \left(0.5 + (me_{20} + 0.5) \cdot 2^{cb_{10}}\right) \cdot \mathrm{ULP}(ir_{22}) < QfltLowerLimit \\ 0, & ir_{22} - \left(0.5 + (me_{20} + 0.5) \cdot 2^{cb_{10}}\right) \cdot \mathrm{ULP}(ir_{22}) > QfltUpperLimit \\ 0.5 + (me_{20} + 0.5) \cdot 2^{cb_{10}}, & \text{otherwise} \end{cases}$$

,

$$ir_{23} = \begin{cases} DfltLowerLimit, & ir_{21} < DfltLowerLimit \\ DfltUpperLimit, & ir_{21} > DfltUpperLimit \\ ir_{21}, & \text{otherwise} \end{cases},$$

$$me_{23} = me_{21},$$

$$ir_{24} = \begin{cases} QfltLowerLimit, & ir_{22} < QfltLowerLimit \\ QfltUpperLimit, & ir_{22} > QfltUpperLimit \\ ir_{22}, & \text{otherwise} \end{cases},$$

$$me_{24} = me_{22},$$

$$ir_{25} = ir_{23} / ir_{24},$$

$$me_{25} = 0.5 + \quad 2 \cdot me_{23} + 2 \cdot me_{24},$$

$$ir_{26} = fltVelocity \cdot fltWIGain,$$

$$me_{26} = 0.5,$$

$$ir_{27} = ir_{26} \cdot ir_{12},$$

$$me_{27} = 0.5 + 2 \cdot me_{26} + 2 \cdot me_{12},$$

$$ir_{28} = -ir_{26} \cdot ir_{9},$$

$$me_{28} = 0.5 + 2 \cdot me_{26} + 2 \cdot me_{9},$$

$$ir_{29} = ir_{23} \cdot fltEGain,$$

$$me_{29} = 2 \cdot me_{23},$$

$$ir_{30} = ir_{24} \cdot fltEGain,$$

$$me_{30} = 2 \cdot me_{24},$$

$$ie_{21} = \max(\mathrm{fe}(|ir_{27}| + me_{27} \cdot \mathrm{ULP}(ir_{27})), \mathrm{fe}(|ir_{29}| + me_{29} \cdot \mathrm{ULP}(ir_{29}))),$$

$$ir_{31} = ir_{27} + ir_{29},$$

$$ie_{22} = \mathrm{fe}(ir_{31}),$$

$$cb_{11} = \begin{cases} 0, & ie_{21} - ie_{22} \le 0 \\ ie_{21} - ie_{22}, & ie_{21} - ie_{22} > 0 \end{cases},$$

$$me_{31} = 0.5 + (me_{27} + me_{29}) \cdot 2^{cb_{11}},$$

$$ie_{23} = \max(\mathrm{fe}(|ir_{28}| + me_{28} \cdot \mathrm{ULP}(ir_{28})), \mathrm{fe}(|ir_{30}| + me_{30} \cdot \mathrm{ULP}(ir_{30}))),$$

$$ir_{32} = ir_{28} + ir_{30},$$

$$ie_{24} = \mathrm{fe}(ir_{32}),$$

$$cb_{12} = \begin{cases} 0, & ie_{23} - ie_{24} \le 0 \\ ie_{23} - ie_{24}, & ie_{23} - ie_{24} > 0 \end{cases},$$

$$me_{32} = 0.5 + (me_{28} + me_{30}) \cdot 2^{cb_{12}},$$

$$ir_{33} = ir_{3} \cdot fltUGain,$$

$$me_{33} = 2 \cdot me_{3},$$

$$ir_{34} = ir_{6} \cdot fltUGain,$$

$$me_{34} = 2 \cdot me_{6},$$

$$ie_{25} = \max(\,\mathrm{fe}(\,|ir_{31}| + me_{31} \cdot \mathrm{ULP}(\,ir_{31}\,)\,),\, \mathrm{fe}(\,|ir_{33}| + me_{33} \cdot \mathrm{ULP}(\,ir_{33}\,)\,)\,),$$

$$ir_{35} = ir_{31} + ir_{33},$$

$$ie_{26} = \mathrm{fe}(\,ir_{35}\,),$$

$$cb_{13} = \begin{cases} 0, & ie_{25} - ie_{26} \le 0 \\ ie_{25} - ie_{26}, & ie_{25} - ie_{26} > 0 \end{cases},$$

$$me_{35} = 0.5 + (\,me_{31} + me_{33}\,) \cdot 2^{cb_{13}},$$

$$ie_{27} = \max(\,\mathrm{fe}(\,|ir_{32}| + me_{32} \cdot \mathrm{ULP}(\,ir_{32}\,)\,),\, \mathrm{fe}(\,|ir_{34}| + me_{34} \cdot \mathrm{ULP}(\,ir_{34}\,)\,)\,),$$

$$ir_{36} = ir_{32} + ir_{34},$$

$$ie_{28} = \mathrm{fe}(\,ir_{36}\,),$$

$$cb_{13} = \begin{cases} 0, & ie_{27} - ie_{28} \le 0 \\ ie_{27} - ie_{28}, & ie_{27} - ie_{28} > 0 \end{cases},$$

$$me_{36} = 0.5 + (\,me_{32} + me_{34}\,) \cdot 2^{cb_{13}},$$

$$ie_{29} = \max(\,\mathrm{fe}(\,|ir_{35}| + me_{35} \cdot \mathrm{ULP}(\,ir_{35}\,)\,),\, \mathrm{fe}(\,|pIObsrvin\_1Arg1| + 0.5 \cdot \mathrm{ULP}(\,pIObsrvin\_1Arg1\,)\,)\,),$$

$$ir_{37} = ir_{35} + pIObsrvin\_1Arg1,$$

$$ie_{30} = \mathrm{fe}(\,ir_{37}\,),$$

$$cb_{14} = \begin{cases} 0, & ie_{29} - ie_{30} \le 0 \\ ie_{29} - ie_{30}, & ie_{29} - ie_{30} > 0 \end{cases},$$

$$me_{37} = 0.5 + (\,me_{35} + 0.5\,) \cdot 2^{cb_{14}},$$

$$ie_{31} = \max(\,\mathrm{fe}(\,|ir_{36}| + me_{36} \cdot \mathrm{ULP}(\,ir_{36}\,)\,),\, \mathrm{fe}(\,|pIObsrvin\_1Arg2| + 0.5 \cdot \mathrm{ULP}(\,pIObsrvin\_1Arg2\,)\,)\,),$$

$$ir_{38} = ir_{36} + pIObsrvin\_1Arg2,$$

$$ie_{32} = \mathrm{fe}(\,ir_{38}\,),$$

$$cb_{15} = \begin{cases} 0, & ie_{31} - ie_{32} \le 0 \\ ie_{31} - ie_{32}, & ie_{31} - ie_{32} > 0 \end{cases},$$

$$me_{38} = 0.5 + (\,me_{36} + 0.5\,) \cdot 2^{cb_{15}},$$

$$ir_{39} = pIObsrvArg1 \cdot fltIGain,$$

$me_{39} = 1$,

$ir_{40} = pIObsrvArg2 \cdot fltIGain$,

$me_{40} = 1$,

$ie_{33} = \max(\mathrm{fe}(|ir_{37}| + me_{37} \cdot \mathrm{ULP}(ir_{37})), \mathrm{fe}(|ir_{39}| + \mathrm{ULP}(ir_{39})))$,

$ir_{41} = ir_{37} + ir_{39}$,

$ie_{34} = \mathrm{fe}(ir_{41})$,

$$cb_{16} = \begin{cases} 0, & ie_{33} - ie_{34} \le 0 \\ ie_{33} - ie_{34}, & ie_{33} - ie_{34} > 0 \end{cases},$$

$me_{41} = 0.5 + (me_{37} + 1) \cdot 2^{cb_{16}}$,

$ie_{35} = \max(\mathrm{fe}(|ir_{38}| + me_{38} \cdot \mathrm{ULP}(ir_{38})), \mathrm{fe}(|ir_{40}| + \mathrm{ULP}(ir_{40})))$,

$ir_{42} = ir_{38} + ir_{40}$,

$ie_{36} = \mathrm{fe}(ir_{42})$,

$$cb_{17} = \begin{cases} 0, & ie_{35} - ie_{36} \le 0 \\ ie_{35} - ie_{36}, & ie_{35} - ie_{36} > 0 \end{cases},$$

$me_{42} = 0.5 + (me_{38} + 1) \cdot 2^{cb_{17}}$,

$me_{43} = 6 + me_{25}$,

then

Table 3-2. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pParamD.fltInErrK1 output

Subset of input domain	Worst-case error bounds for pParamD.fltInErrK1 output [ulp]	Allowed specific values (regardless the error bounds) for pParamD.fltInErrK1 output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e.	$\langle -me_{13}, me_{13} \rangle$	N/A

Table continues on the next page...

Table 3-2. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pParamD.fltInErrK1 output (continued)

Subset of input domain	Worst-case error bounds for pParamD.fltInErrK1 output [ulp]	Allowed specific values (regardless the error bounds) for pParamD.fltInErrK1 output
$\left\{X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$		
Normalized or zero input values which cause an underflow, i.e. $\left\{X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n: ir_n \ne \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n: ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-3. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pParamQ.fltInErrK1 output

Subset of input domain	Worst-case error bounds for pParamQ.fltInErrK1 output [ulp]	Allowed specific values (regardless the error bounds) for pParamQ.fltInErrK1 output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \ne \varnothing \vee X \cap M \ne \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$	$\langle -me_{14}, me_{14} \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n: ir_n \ne \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n: ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-4. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pParamD.fltAcc output

Subset of input domain	Worst-case error bounds for pParamD.fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for pParamD.fltAcc output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$	$\langle -me_{21}, me_{21} \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \exists n : ir_n \neq \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-5. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pParamQ.fltAcc output

Subset of input domain	Worst-case error bounds for pParamQ.fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for pParamQ.fltAcc output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$	$\langle -me_{22}, me_{22} \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \exists n : ir_n \neq \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table continues on the next page...

Table 3-5. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pParamQ.fltAcc output (continued)

Subset of input domain	Worst-case error bounds for pParamQ.fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for pParamQ.fltAcc output
Normalized or zero input values which cause an overflow, i.e. $\left\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \; \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-6. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pPhaseErr output

Subset of input domain	Worst-case error bounds for pPhaseErr output [ulp]	Allowed specific values (regardless the error bounds) for pPhaseErr output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{ X \mid X \cap D \ne \varnothing \vee X \cap M \ne \varnothing \}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi \}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \; \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \right\}$	$\langle -me_{22}, me_{22} \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \; \exists n : ir_n \ne \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \; \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results with pEObsrv.fltArg2 close to zero, i.e. $\left\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \; \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42}, \; \lvert ir_{22} \rvert \le me_{22} \cdot \mathrm{ULP}(ir_{22}) \right\}$	$(-Inf, Inf)$	N/A

Table 3-7. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pIObsrvIn_1.fltArg1 output

Subset of input domain	Worst-case error bounds for pIObsrvIn_1.fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for pIObsrvIn_1.fltArg1 output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert \leq \pi, \ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42}\right\}$	$\langle -me_{35}, me_{35}\rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert \leq \pi, \ \exists n: ir_n \neq \varnothing \wedge \lvert ir_n\rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert \leq \pi, \ \exists n: ir_n \in M \vee \lvert ir_n\rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-8. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pIObsrvIn_1.fltArg2 output

Subset of input domain	Worst-case error bounds for pIObsrvIn_1.fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for pIObsrvIn_1.fltArg2 output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert \leq \pi, \ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42}\right\}$	$\langle -me_{36}, me_{36}\rangle$	N/A

Table continues on the next page...

Table 3-8. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pIObsrvIn_1.fltArg2 output (continued)

Subset of input domain	Worst-case error bounds for pIObsrvIn_1.fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for pIObsrvIn_1.fltArg2 output
Normalized or zero input values which cause an underflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \ne \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-9. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pIObsrv.fltArg1 output

Subset of input domain	Worst-case error bounds for pIObsrv.fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for pIObsrv.fltArg1 output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{ X \mid X \cap D \ne \varnothing \vee X \cap M \ne \varnothing \}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi \}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{array} \right\}$	$\langle -me_{41}, me_{41} \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \ne \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-10. AMCLIB_BemfObsrvDQ_FLT Worst-Case Error Bounds - pIObsrv.fltArg2 output

Subset of input domain	Worst-case error bounds for pIObsrv.fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for pIObsrv.fltArg2 output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \emptyset \vee X \cap M \neq \emptyset\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$	$\langle -me_{42}, me_{42} \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \exists n: ir_n \neq \emptyset \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \exists n: ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

3.2 Function AMCLIB_TrackObsrv_FLT

3.2.1 Declaration

```
void AMCLIB_TrackObsrv_FLT(tFloat fltPhaseErr, tFloat *pPosEst, tFloat *pVelocityEst,
AMCLIB_TRACK_OBSRV_T_FLT *pCtrl);
```

3.2.2 Arguments

Table 3-11. AMCLIB_TrackObsrv_FLT arguments

Type	Name	Direction	Description
tFloat	fltPhaseErr	**input**	Input signal representing phase error of system to be estimated.
tFloat *	pPosEst	**output**	Estimated output position.
tFloat *	pVelocityEst	**output**	Estimated output velocity.

Table continues on the next page...

Table 3-11. AMCLIB_TrackObsrv_FLT arguments (continued)

Type	Name	Direction	Description
AMCLIB_TRACK_OBSRV_T_FLT *	pCtrl	**input, output**	Pointer to a tracking observer structure AMCLIB_TRACK_OBSRV_T_FLT, which contains algorithm coefficients.

3.2.3 Worst-Case Error Bounds

Let $(fltPhaseErr, fltCC1sc, fltCC2sc, fltAcc, fltInErrK1, fltUpperLimit, fltLowerLimit, fltState, fltInK1, fltC1) \in X$ be a vector of inputs to AMCLIB_TrackObsrv_FLT,

$fltAcc, fltState, fltInK1$ inputs represent real numbers with an error of max. ±0.5 ulp,

$$ir_1 = fltPhaseErr \cdot fltCC1sc,$$

$$me_1 = 0,$$

$$ir_2 = fltInErrK1 \cdot fltCC2sc,$$

$$me_2 = 0,$$

$$ir_3 = ir_1 + ir_2,$$

$$me_3 = 1,$$

$$ie_1 = \max(\mathrm{fe}(|ir_3| + me_3 \cdot \mathrm{ULP}(ir_3)), \mathrm{fe}(|fltAcc| + 0.5 \cdot \mathrm{ULP}(fltAcc))),$$

$$ir_4 = \begin{cases} fltLowerLimit0, & ir_3 + fltAcc < fltLowerLimit \\ fltUpperLimit, & ir_3 + fltAcc > fltUpperLimit \\ ir_3 + fltAcc, & \text{otherwise} \end{cases},$$

$$ie_2 = \mathrm{fe}(ir_4),$$

$$cb_1 = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$$me_4 = \begin{cases} 0, & ir_4 + \left(0.5 + 1.5 \cdot 2^{cb_1}\right) \cdot \mathrm{ULP}(ir_4) < fltLowerLimit \\ 0, & ir_4 - \left(0.5 + 1.5 \cdot 2^{cb_1}\right) \cdot \mathrm{ULP}(ir_4) > fltUpperLimit \\ 0.5 + 1.5 \cdot 2^{cb_1}, & \text{otherwise} \end{cases},$$

$$ie_3 = \max(\mathrm{fe}(|ir_4| + me_4 \cdot \mathrm{ULP}(ir_4)), \mathrm{fe}(|fltInK1| + 0.5 \cdot \mathrm{ULP}(fltInK1))),$$

$$ir_5 = ir_4 + fltInK1,$$

$$ie_4 = \mathrm{fe}(ir_5),$$

$$cb_2 = \begin{cases} 0, & ie_3 - ie_4 \le 0 \\ ie_3 - ie_4, & ie_3 - ie_4 > 0 \end{cases},$$

$$me_5 = 0.5 + (me_4 + 0.5) \cdot 2^{cb_2},$$

$$ir_6 = ir_5 \cdot fltC1,$$

$$me_6 = 2 \cdot me_5,$$

$$ie_5 = \max(\mathrm{fe}(|ir_6| + me_6 \cdot \mathrm{ULP}(ir_6)), \mathrm{fe}(|fltState| + 0.5 \cdot \mathrm{ULP}(fltState))),$$

$$ir_7 = ir_6 + fltState,$$

$$ie_6 = \mathrm{fe}(ir_7),$$

$$cb_3 = \begin{cases} 0, & ie_5 - ie_6 \le 0 \\ ie_5 - ie_6, & ie_5 - ie_6 > 0 \end{cases},$$

$$me_7 = 0.5 + (me_6 + 0.5) \cdot 2^{cb_3},$$

then

Table 3-12. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pVelocityEst output

Subset of input domain	Worst-case error bounds for pVelocityEst output [ulp]	Allowed specific values (regardless the error bounds) for pVelocityEst output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert \le \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$	$\langle -me_4, me_4 \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert \le \pi, \\ \exists n : ir_n \neq \varnothing \wedge \lvert ir_n\rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr\rvert \le \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n\rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-13. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pParamPI.fltAcc output

Subset of input domain	Worst-case error bounds for pParamPI.fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for pParamPI.fltAcc output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$	$\langle -me_4, me_4 \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \exists n : ir_n \neq \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \leq \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-14. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pParamPI.fltInErrK1 output

Subset of input domain	Worst-case error bounds for pParamPI.fltInErrK1 output [ulp]	Allowed specific values (regardless the error bounds) for pParamPI.fltInErrK1 output
Entire input domain	0	N/A

Table 3-15. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pPosEst output

Subset of input domain	Worst-case error bounds for pPosEst output [ulp]	Allowed specific values (regardless the error bounds) for pPosEst output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table continues on the next page...

Table 3-15. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pPosEst output (continued)

Subset of input domain	Worst-case error bounds for pPosEst output [ulp]	Allowed specific values (regardless the error bounds) for pPosEst output
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{matrix}\right\}$	$\langle -me_7, me_7 \rangle$	N/A
Normalized or zero input values which lead to normalized or zero results with output wrap-around, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ ir \in (N \cup \{-0, +0\})^{42} \wedge \\ \left(\begin{matrix} \lvert ir_7 \rvert + me_7 \cdot \mathrm{ULP}(ir_7) > \pi \vee \\ \lvert ir_7 \rvert - me_7 \cdot \mathrm{ULP}(ir_7) < -\pi \end{matrix}\right) \end{matrix}\right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \neq \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{\begin{matrix} X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{matrix}\right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-16. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pParamInteg.fltState output

Subset of input domain	Worst-case error bounds for pParamInteg.fltState output [ulp]	Allowed specific values (regardless the error bounds) for pParamInteg.fltState output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{X \mid X \cap D \neq \varnothing \vee X \cap M \neq \varnothing\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e.	$\langle -me_7, me_7 \rangle$	N/A

Table continues on the next page...

Table 3-16. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pParamInteg.fltState output (continued)

Subset of input domain	Worst-case error bounds for pParamInteg.fltState output [ulp]	Allowed specific values (regardless the error bounds) for pParamInteg.fltState output
$\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{array} \right\}$		
Normalized or zero input values which cause an underflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \ne \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

Table 3-17. AMCLIB_TrackObsrv_FLT Worst-Case Error Bounds - pParamInteg.fltInK1 output

Subset of input domain	Worst-case error bounds for pParamInteg.fltInK1 output [ulp]	Allowed specific values (regardless the error bounds) for pParamInteg.fltInK1 output
Any of the inputs is denormalized, NaN, or infinity, i.e. $\{ X \mid X \cap D \ne \varnothing \vee X \cap M \ne \varnothing \}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values with pPhaseErr beyond the allowed range, i.e. $\{ X \mid X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert > \pi \}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \boldsymbol{ir} \in (N \cup \{-0, +0\})^{42} \end{array} \right\}$	$\langle -me_4, me_4 \rangle$	N/A
Normalized or zero input values which cause an underflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \ne \varnothing \wedge \lvert ir_n \rvert - me_n \cdot \mathrm{ULP}(ir_n) \in D \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause an overflow, i.e. $\left\{ X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\} \wedge \lvert pPhaseErr \rvert \le \pi, \\ \exists n : ir_n \in M \vee \lvert ir_n \rvert + me_n \cdot \mathrm{ULP}(ir_n) = Inf \end{array} \right\}$	$(-Inf, Inf)$	Arbitrary floating-point value (including Inf, -Inf, NaN)

3.3 Function GDFLIB_FilterFIR_FLT

3.3.1 Declaration

```
tFloat GDFLIB_FilterFIR_FLT(tFloat fltIn, const GDFLIB_FILTERFIR_PARAM_T_FLT *const pParam,
GDFLIB_FILTERFIR_STATE_T_FLT *const pState);
```

3.3.2 Arguments

Table 3-18. GDFLIB_FilterFIR_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input value.
const GDFLIB_FILTERFIR_PARAM_T_FLT *const	pParam	**input**	Pointer to a parameter structure.
GDFLIB_FILTERFIR_STATE_T_FLT *const	pState	**input, output**	Pointer to a filter state structure.

3.3.3 Worst-Case Error Bounds

Let $(fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X$ be a vector of inputs to GDFLIB_FilterFIR_FLT,

$\boldsymbol{CoefBuf} = \left(CoefBuf_0, CoefBuf_1, \ldots, CoefBuf_{u32Order}\right)$,

$\boldsymbol{InBuf} = \left(fltIn_t, fltIn_{t-1}, \ldots, fltIn_{t-u32Order}\right)$, where t is the current time,

$refResult$ be the theoretical exact result,

$\boldsymbol{mr} = \boldsymbol{CoefBuf} \odot \boldsymbol{InBuf}$, i.e. a vector of element-wise multiplication results of vectors $\boldsymbol{CoefBuf}$ and $\boldsymbol{InBuf}$,

$\boldsymbol{sr} = \left(sr_1, sr_2, \ldots, sr_{u32Order-1}\right), sr_k = \sum_{n=0}^{k} CoefBuf_n \cdot InBuf_n$, i.e. a vector of partial sums of element-wise multiplication results of vectors $\boldsymbol{CoefBuf}$ and $\boldsymbol{InBuf}$,

$ie_1 = \max_n\left(\mathrm{fe}\left(mr_n\right)\right)$,

$ie_2 = \mathrm{fe}\left(refResult\right)$,

$$cb_1 = \begin{cases} 0, & ie_1 - ie_2 \leq 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$$me = 0.5 \cdot (u32Order + 1) \cdot 2^{cb_1} + \sum_{n=1}^{\mathrm{ceil}\left(\frac{u32Order-2}{2}\right)} \mathrm{ceil}\left(\log_2(n+1)\right) + \sum_{n=1}^{\mathrm{floor}\left(\frac{u32Order-2}{2}\right)} \mathrm{ceil}\left(\log_2(n+1)\right),$$

then

Table 3-19. GDFLIB_FilterFIR_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e. $\left\{ (fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X \;\middle|\; X \cap M = \varnothing \wedge X \cap D \neq \varnothing \right\}$	$(-Inf, Inf)$	N/A
Input values which contain at least one NaN, Inf, or -Inf, i.e. $\{ (fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X \mid X \cap M \neq \varnothing \}$	N/A	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{array}{c} (fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ mr \in (N \cup \{-0, +0\})^{u32Order+1}, \\ sr \in (N \cup \{-0, +0\})^{u32Order-1}, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{array}{c} (fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \in M \end{array} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{array}{c} (fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate result overflow, i.e.	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table continues on the next page...

Table 3-19. GDFLIB_FilterFIR_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} (fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \exists\, mr_n \in M \vee \\ \exists\, (\lvert sr_n \rvert + me \cdot \mathrm{ULP}(sr_n)) \in M \end{array} \right\}$		
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{array}{c} (fltIn, \boldsymbol{CoefBuf}, \boldsymbol{InBuf}) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \exists\, mr_n \in D \vee \\ \exists\, (\lvert sr_n \rvert - me \cdot \mathrm{ULP}(sr_n)) \in D, sr \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	N/A

3.4 Function GDFLIB_FilterIIR1_FLT

3.4.1 Declaration

```
tFloat GDFLIB_FilterIIR1_FLT(tFloat fltIn, GDFLIB_FILTER_IIR1_T_FLT *const pParam);
```

3.4.2 Arguments

Table 3-20. GDFLIB_FilterIIR1_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Value of the input signal to be filtered in step (k). Input is a 32-bit number that contains a single precision floating point value.
GDFLIB_FILTER_IIR1_T_FLT *const	pParam	**input, output**	Pointer to a filter structure with a filter buffer and filter parameters. Arguments of the structure contain single precision floating point values.

3.4.3 Worst-Case Error Bounds

Let $(fltIn, fltB0, fltB1, flA1, fltFiltBufferX, fltFiltBufferY) \in X$ be a vector of inputs to GDFLIB_FilterIIR1_FLT,

$refResult$ be the theoretical exact result,

$\boldsymbol{mr} = (fltIn, fltFiltBufferX, fltFiltBufferY) \odot (fltB0, fltB1, fltA1)$, i.e. a vector of element-wise multiplication results of vectors $(fltIn, fltFiltBufferX, fltFiltBufferY)$ and $(fltB0, fltB1, fltA1)$,

$sr = fltIn \cdot fltB0 + fltFiltBufferX \cdot fltB1$,

$ie_1 = \max_n(\mathrm{fe}(mr_n))$,

$ie_2 = \mathrm{fe}(refResult)$,

$cb = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases}$,

$me = 0.5 + 2.5 \cdot 2^{cb}$,

the $fltFiltBufferY$ input represents a real number with an error of max. ±0.5 ulp,

then

Table 3-21. GDFLIB_FilterIIR1_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \middle\| X \cap M = \varnothing \wedge X \cap D \neq \varnothing \right\}$	$(-Inf, Inf)$	N/A
Input values which contain at least one NaN, Inf, or –Inf, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \middle\| X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e.	$\langle -me, me \rangle$	N/A

Table continues on the next page...

Table 3-21. GDFLIB_FilterIIR1_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\vert\, \begin{matrix} X \subset N \cup \{-0, +0\}, \boldsymbol{mr} \in N^{u32Order+1}, \\ sr \in N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$		
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\vert\, \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\vert\, \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate result overflow with normalized or zero reference output, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\vert\, \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \boldsymbol{mr} \in M^{u32Order+1} \vee \\ \lvert sr \rvert + me \cdot \mathrm{ULP}(sr) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause intermediate result underflow with normalized or zero reference output, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\vert\, \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \boldsymbol{mr} \in D^{u32Order+1} \vee \\ \lvert sr \rvert - me \cdot \mathrm{ULP}(sr) \in D, sr \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table 3-22. GDFLIB_FilterIIR1_FLT Worst-Case Error Bounds - fltFiltBufferY Output

Subset of input domain	Worst-case error bounds for fltFiltBufferY output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferY output
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\|\, X \cap M = \varnothing \wedge X \cap D \neq \varnothing \right\}$	$(-Inf, Inf)$	N/A
Input values which contain at least one NaN, Inf, or –Inf, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\|\, X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\|\, \begin{matrix} X \subset N \cup \{-0, +0\}, \boldsymbol{mr} \in N^{u32Order+1}, \\ sr \in N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\|\, \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \,\middle\|\, \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate result overflow with normalized or zero reference output, i.e.	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table continues on the next page...

Table 3-22. GDFLIB_FilterIIR1_FLT Worst-Case Error Bounds - fltFiltBufferY Output (continued)

Subset of input domain	Worst-case error bounds for fltFiltBufferY output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferY output
$\left\{ \begin{array}{c} \left(\begin{array}{c} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{array} \right) \in X \left\| \right. \\ X \subset N \cup \{-0, +0\}, \\ \boldsymbol{mr} \in M^{u32Order+1} \vee \\ \lvert sr \rvert + me \cdot \mathrm{ULP}(sr) \in M \end{array} \right\}$		
Normalized or zero input values which cause intermediate result underflow with normalized or zero reference output, i.e. $\left\{ \begin{array}{c} \left(\begin{array}{c} fltIn, fltB0, fltB1, fltA1, \\ fltFiltBufferX, fltFiltBufferY \end{array} \right) \in X \left\| \right. \\ X \subset N \cup \{-0, +0\}, \\ \boldsymbol{mr} \in D^{u32Order+1} \vee \\ \lvert sr \rvert - me \cdot \mathrm{ULP}(sr) \in D, sr \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table 3-23. GDFLIB_FilterIIR1_FLT Worst-Case Error Bounds - fltFiltBufferX Output

Subset of input domain	Worst-case error bounds for fltFiltBufferX output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferX output
Entire input domain	0	*fltIn*

3.5 Function GDFLIB_FilterIIR2_FLT

3.5.1 Declaration

```
tFloat GDFLIB_FilterIIR2_FLT(tFloat fltIn, GDFLIB_FILTER_IIR2_T_FLT *const pParam);
```

3.5.2 Arguments

Table 3-24. GDFLIB_FilterIIR2_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Value of the input signal to be filtered in step (k). Input is a 32-bit number that contains a single precision floating point value.
GDFLIB_FILTER_IIR2_T_FLT *const	pParam	**input, output**	Pointer to a filter structure with a filter buffer and filter parameters. Arguments of the structure contain single precision floating point values.

3.5.3 Worst-Case Error Bounds

Let $(fltIn, fltB0, fltB1, fltB2, flA1, flA2, \boldsymbol{fltFiltBufferX}, \boldsymbol{fltFiltBufferY}) \in X$ be a vector of inputs to GDFLIB_FilterIIR2_FLT,

$\boldsymbol{fltFiltBufferX} = \left(fltFiltBufferX_0, fltFiltBufferX_1\right)$,

$\boldsymbol{fltFiltBufferY} = \left(fltFiltBufferY_0, fltFiltBufferY_1\right)$,

$refResult$ be the theoretical exact result,

$\boldsymbol{mr} = \left(fltIn, fltFiltBufferX, fltFiltBufferY\right) \odot \left(fltB0, fltB1, fltB2, fltA1, fltA2\right)$, i.e. a vector of element-wise multiplication results of vectors $(fltIn, \boldsymbol{fltFiltBufferX}, \boldsymbol{fltFiltBufferY})$ and $(fltB0, fltB1, fltB2, flA1, flA2)$,

$sr_1 = fltIn \cdot fltB_0 + fltFiltBufferX_0 \cdot fltB1$,

$sr_2 = sr_1 + fltFiltBufferX_1 \cdot fltB2$,

$sr_3 = sr_2 + fltFiltBufferY_0 \cdot fltA1$,

$\boldsymbol{sr} = \left(sr_1, sr_2, sr_3\right)$,

$ie_1 = \max_n\left(\mathrm{fe}\left(mr_n\right)\right)$,

$ie_2 = \mathrm{fe}\left(refResult\right)$,

$$cb = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me = 0.5 + 5 \cdot 2^{cb}$,

the $fltFiltBufferY_0$ input represents a real number with an error of max. ±0.5 ulp,

then

Table 3-25. GDFLIB_FilterIIR2_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \;\middle\|\; X \cap M = \varnothing \wedge X \cap D \neq \varnothing \right\}$	$(-Inf, Inf)$	N/A
Input values which contain at least one NaN, Inf, or –Inf, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \;\middle\|\; X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \;\middle\|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \boldsymbol{mr} \in (N \cup \{-0, +0\})^{u32Order+1}, \\ \boldsymbol{sr} \in (N \cup \{-0, +0\})^{u32Order-1}, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \;\middle\|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \;\middle\|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate result overflow, i.e.	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table continues on the next page...

Table 3-25. GDFLIB_FilterIIR2_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \middle\| \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \exists\, mr_n \in M \vee \\ \exists\, (\lvert sr_n \rvert + me \cdot \mathrm{ULP}(sr_n)) \in M \end{array} \right\}$		
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \middle\| \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \exists\, mr_n \in D \vee \\ \exists\, (\lvert sr_n \rvert - me \cdot \mathrm{ULP}(sr_n)) \in D, sr \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table 3-26. GDFLIB_FilterIIR2_FLT Worst-Case Error Bounds - fltFiltBufferY$_0$ Output

Subset of input domain	Worst-case error bounds for fltFiltBufferY$_0$ output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferY$_0$ output
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \middle\| X \cap M = \varnothing \wedge X \cap D \neq \varnothing \right\}$	$(-Inf, Inf)$	N/A
Input values which contain at least one NaN, Inf, or –Inf, i.e. $\left\{ \begin{pmatrix} fltIn, fltB0, fltB1, fltB2, fltA1, fltA2, \\ fltFiltBufferX, fltFiltBufferY \end{pmatrix} \in X \middle\| X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e.	$\langle -me, me \rangle$	N/A

Table continues on the next page...

Table 3-26. GDFLIB_FilterIIR2_FLT Worst-Case Error Bounds - fltFiltBufferY$_0$ Output (continued)

Subset of input domain	Worst-case error bounds for fltFiltBufferY$_0$ output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferY$_0$ output
$\left\{ \begin{matrix} \begin{pmatrix} fltIn,\ fltB0,\ fltB1,\ fltB2,\ fltA1,\ fltA2, \\ fltFiltBufferX,\ fltFiltBufferY \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \boldsymbol{mr} \in (N \cup \{-0,\ +0\})^{u32Order+1}, \\ \boldsymbol{sr} \in (N \cup \{-0,\ +0\})^{u32Order-1}, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$		
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{matrix} \begin{pmatrix} fltIn,\ fltB0,\ fltB1,\ fltB2,\ fltA1,\ fltA2, \\ fltFiltBufferX,\ fltFiltBufferY \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{matrix} \begin{pmatrix} fltIn,\ fltB0,\ fltB1,\ fltB2,\ fltA1,\ fltA2, \\ fltFiltBufferX,\ fltFiltBufferY \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \begin{matrix} \begin{pmatrix} fltIn,\ fltB0,\ fltB1,\ fltB2,\ fltA1,\ fltA2, \\ fltFiltBufferX,\ fltFiltBufferY \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \exists\ mr_n \in M \vee \\ \exists\ (\lvert sr_n \rvert + me \cdot \mathrm{ULP}(sr_n)) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{matrix} \begin{pmatrix} fltIn,\ fltB0,\ fltB1,\ fltB2,\ fltA1,\ fltA2, \\ fltFiltBufferX,\ fltFiltBufferY \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \exists\ mr_n \in D \vee \\ \exists\ (\lvert sr_n \rvert - me \cdot \mathrm{ULP}(sr_n)) \in D,\ sr \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table 3-27. GDFLIB_FilterIIR2_FLT Worst-Case Error Bounds - fltFiltBufferY$_1$ Output

Subset of input domain	Worst-case error bounds for fltFiltBufferY$_1$ output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferY$_1$ output
Entire input domain	0	*fltFiltBufferY*$_0$

Table 3-28. GDFLIB_FilterIIR2_FLT Worst-Case Error Bounds - fltFiltBufferX$_0$ Output

Subset of input domain	Worst-case error bounds for fltFiltBufferX$_0$ output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferX$_0$ output
Entire input domain	0	*fltIn*

Table 3-29. GDFLIB_FilterIIR2_FLT Worst-Case Error Bounds - fltFiltBufferX$_1$ Output

Subset of input domain	Worst-case error bounds for fltFiltBufferX$_1$ output [ulp]	Allowed specific values (regardless the error bounds) for fltFiltBufferX$_1$ output
Entire input domain	0	*fltFiltBufferX*$_0$

3.6 Function GDFLIB_FilterMA_FLT

3.6.1 Declaration

```
tFloat GDFLIB_FilterMA_FLT(tFloat fltIn, GDFLIB_FILTER_MA_T_FLT *pParam);
```

3.6.2 Arguments

Table 3-30. GDFLIB_FilterMA_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Value of the input signal to be filtered in step (k). The value is a single precision floating point data type.
GDFLIB_FILTER_MA_T_FLT *	pParam	**input, output**	Pointer to the filter structure with a filter accumulator and a smoothing factor.

3.6.3 Worst-Case Error Bounds

Let $(fltIn, fltAcc, fltLambda) \in X$ be a vector of inputs to GDFLIB_FilterMA_FLT, $refResult$ be the theoretical exact result,

$mr_1 = fltIn \cdot fltLambda$,

$mr_2 = refResult \cdot fltLambda$,

$ie_1 = \max(\mathrm{fe}(mr_1), \mathrm{fe}(fltAcc))$,

$ie_2 = \mathrm{fe}(refResult)$,

$$cb_1 = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me_1 = 0.5 + 0.5 \cdot 2^{cb_1}$,

$refState = refResult - mr_2$,

$ie_3 = \max(\mathrm{fe}(|mr_2| + 2 \cdot me_1 \cdot \mathrm{ULP}(mr_2)), \mathrm{fe}(|refResult| + me_1 \cdot \mathrm{ULP}(refResult)))$,

$ie_4 = \mathrm{fe}(refState)$,

$$cb_2 = \begin{cases} 0, & ie_3 - ie_4 \le 0 \\ ie_3 - ie_4, & ie_3 - ie_4 > 0 \end{cases},$$

$me_2 = 0.5 + me_1 + (2 \cdot me_1) \cdot 2^{cb_2}$,

then

Table 3-31. GDFLIB_FilterMA_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e. $\left\{ (fltIn, fltAcc, fltLambda) \in X \;\middle|\; X \cap M = \varnothing \wedge X \cap D \neq \varnothing \right\}$	$(-Inf, Inf)$	N/A
Input values which contain at least one NaN, Inf, or –Inf, i.e. $\{ (fltIn, fltAcc, fltLambda) \in X \mid X \cap M \neq \varnothing \}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero result, i.e.	$\langle -me_1, me_1 \rangle$	N/A

Table continues on the next page...

Table 3-31. GDFLIB_FilterMA_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{\begin{matrix} (fltIn,\ fltAcc,\ fltLambda) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ mr_1 \in N \cup \{-0,\ +0\}, \\ refResult = 0 \vee \\ \left(\begin{matrix} \lvert refResult \rvert - me_1 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me_1 \cdot \mathrm{ULP}(refResult) \neq Inf \end{matrix}\right), \end{matrix}\right\}$		
Normalized or zero input values which cause result overflow, i.e. $\left\{\begin{matrix} (fltIn,\ fltAcc,\ fltLambda) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert + me_1 \cdot \mathrm{ULP}(refResult) \in M \end{matrix}\right\}$	$\langle -Inf, Inf \rangle$	{Inf, -Inf, NaN}
Normalized or zero input values which cause result underflow, i.e. $\left\{\begin{matrix} (fltIn,\ fltAcc,\ fltLambda) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me_1 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix}\right\}$	$\langle -Inf, Inf \rangle$	N/A
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{\begin{matrix} (fltIn,\ fltAcc,\ fltLambda) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ mr_1 \in \{Inf,\ -Inf\} \end{matrix}\right\}$	$\langle -Inf, Inf \rangle$	{Inf, -Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{\begin{matrix} (fltIn,\ fltAcc,\ fltLambda) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ mr_1 \in D \end{matrix}\right\}$	$\langle -Inf, Inf \rangle$	N/A

Table 3-32. GDFLIB_FilterMA_FLT Worst-Case Error Bounds - fltAcc Output

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e.	$\langle -Inf, Inf \rangle$	N/A

Table continues on the next page...

Table 3-32. GDFLIB_FilterMA_FLT Worst-Case Error Bounds - fltAcc Output (continued)

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
$\left\{ \begin{matrix} (fltIn, fltAcc, fltLambda) \in X \mid \\ X \cap M = \varnothing \wedge X \cap D \neq \varnothing \end{matrix} \right\}$		
Input values which contain at least one NaN, Inf, or –Inf, i.e. $\{ (fltIn, fltAcc, fltLambda) \in X \mid X \cap M \neq \varnothing \}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero fltAcc output, i.e. $\left\{ \begin{matrix} (fltIn, fltAcc, fltLambda) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ mr_1 \in N \cup \{-0, +0\}, mr_2 \in N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \left(\begin{matrix} \lvert refResult \rvert - me_1 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me_1 \cdot \mathrm{ULP}(refResult) \neq Inf \end{matrix} \right), \\ refState = 0 \vee \\ \left(\begin{matrix} \lvert refState \rvert - me_2 \cdot \mathrm{ULP}(refState) \notin D \wedge \\ \lvert refState \rvert + me_2 \cdot \mathrm{ULP}(refState) \neq Inf \end{matrix} \right) \end{matrix} \right\}$	$\langle -me_2, me_2 \rangle$	N/A
Normalized or zero input values which cause *fltAcc* output overflow, i.e. $\left\{ \begin{matrix} (fltIn, fltAcc, fltLambda) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \lvert refState \rvert + me_2 \cdot \mathrm{ULP}(refState) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Normalized or zero input values which cause *fltAcc* output underflow, i.e. $\left\{ \begin{matrix} (fltIn, fltAcc, fltLambda) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \lvert refState \rvert - me_2 \cdot \mathrm{ULP}(refState) \in D \wedge \\ refState \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate results overflow, i.e. $\left\{ \begin{matrix} (fltIn, fltAcc, fltLambda) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ mr_1 \in \{Inf, -Inf\}, \\ \lvert mr_2 \rvert + 2 \cdot me_1 \cdot \mathrm{ULP}(mr_2) = Inf \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table continues on the next page...

Table 3-32. GDFLIB_FilterMA_FLT Worst-Case Error Bounds - fltAcc Output (continued)

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
Normalized or zero input values which cause intermediate results underflow, i.e. $\left\{ \begin{array}{c} (fltIn, fltAcc, fltLambda) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ mr_1 \in D \vee \\ \left(\begin{array}{c} refResult \neq 0 \wedge \\ \lvert refResult \rvert - me_1 \cdot \mathrm{ULP}(refResult) \in D \end{array} \right) \vee \\ (mr_2 \neq 0 \wedge \lvert mr_2 \rvert - 2 \cdot me_1 \cdot \mathrm{ULP}(mr_2) \in D) \end{array} \right\}$	$(-Inf, Inf)$	N/A

3.7 Function GFLIB_Acos_FLT

3.7.1 Declaration

```
tFloat GFLIB_Acos_FLT(tFloat fltIn, const GFLIB_ACOS_T_FLT *const pParam);
```

3.7.2 Arguments

Table 3-33. GFLIB_Acos_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument is a 32-bit number that contains a single precision floating point value.
const GFLIB_ACOS_T_FLT *const	pParam	**input**	Pointer to an array of approximation coefficients. The function alias GFLIB_Acos uses the default coefficients.

3.7.3 Worst-Case Error Bounds

Only the default approximation coefficients are considered.

Let $fltIn \in X$ be an input to GFLIB_Acos_FLT,

then

Table 3-34. GFLIB_Acos_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input NaN or infinity, i.e. $\{ fltIn \in X \mid fltIn \in M \}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Input denormalized, normalized or zero, and within the allowed input interval, i.e. $\left\{ \begin{array}{l} fltIn \in X \mid X \cap M = \emptyset, \\ -1 \leq fltIn \leq 1 \end{array} \right\}$	$\langle -130, 130 \rangle$	N/A
Input normalized and beyond the allowed range, i.e. $\left\{ \begin{array}{l} fltIn \in X \mid X \cap M = \emptyset, \\ fltIn < -1 \vee fltIn > 1 \end{array} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

3.8 Function GFLIB_Asin_FLT

3.8.1 Declaration

```
tFloat GFLIB_Asin_FLT(tFloat fltIn, const GFLIB_ASIN_T_FLT *const pParam);
```

3.8.2 Arguments

Table 3-35. GFLIB_Asin_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument is a 32-bit number that contains a single precision floating point value.
const GFLIB_ASIN_T_FLT *const	pParam	**input**	Pointer to an array of approximation coefficients.

3.8.3 Worst-Case Error Bounds

Only the default approximation coefficients are considered.

Let $fltIn \in X$ be an input to GFLIB_Asin_FLT,

then

Table 3-36. GFLIB_Asin_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input NaN or infinity, i.e. $\{ fltIn \in X \mid fltIn \in M \}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Input denormalized, normalized or zero, and within the allowed input interval, i.e. $\left\{ \begin{matrix} fltIn \in X \mid X \cap M = \varnothing, \\ -1 \le fltIn \le 1 \end{matrix} \right\}$	$\langle -155, 155 \rangle$	N/A
Input normalized and beyond the allowed range, i.e. $\left\{ \begin{matrix} fltIn \in X \mid X \cap M = \varnothing, \\ fltIn < -1 \vee fltIn > 1 \end{matrix} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

3.9 Function GFLIB_Atan_FLT

3.9.1 Declaration

```
tFloat GFLIB_Atan_FLT(tFloat fltIn, const GFLIB_ATAN_T_FLT *const pParam);
```

3.9.2 Arguments

Table 3-37. GFLIB_Atan_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument is a single precision floating point number between (-2^{128}, 2^{128}).
const GFLIB_ATAN_T_FLT *const	pParam	**input**	Pointer to an array of rational polynomial coefficients.

3.9.3 Worst-Case Error Bounds

Only the default approximation coefficients are considered.

Let $fltIn \in X$ be a an input to GFLIB_Atan_FLT,

then

Table 3-38. GFLIB_Atan_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input NaN, i.e. $\{fltIn \in X \mid fltIn = NaN\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Input not a NaN, i.e. $\{fltIn \in X \mid fltIn \neq NaN, \}$	$\langle -3,3 \rangle$	N/A

3.10 Function GFLIB_AtanYX_FLT

3.10.1 Declaration

```
tFloat GFLIB_AtanYX_FLT(tFloat fltInY, tFloat fltInX);
```

3.10.2 Arguments

Table 3-39. GFLIB_AtanYX_FLT arguments

Type	Name	Direction	Description
tFloat	fltInY	**input**	The ordinate of the input vector (y coordinate).
tFloat	fltInX	**input**	The abscissa of the input vector (x coordinate).

3.10.3 Worst-Case Error Bounds

Let $(fltInX, fltInY) \in X$ be a vector of inputs to GFLIB_AtanYX_FLT,

refResult be the theoretical exact result,

then

Table 3-40. GFLIB_AtanYX_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, NaN, or invinity, i.e. $\{ (fltInY, fltInX) \in X \mid X \cap M \neq \varnothing \vee X \cap D \neq \varnothing \}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Both inputs are normalized or zero, results normalized or zero, i.e. $\left\{ \begin{array}{c} (fltInY, fltInX) \in X \mid X \subset N \cup \{-0, +0\}, \\ \frac{fltInY}{fltInX} \in N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - 3.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 3.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -3.5, 3.5 \rangle$	N/A
Both inputs are normalized or zero, result underflow, i.e. $\left\{ \begin{array}{c} (fltInY, fltInX) \in X \mid X \subset N \cup \{-0, +0\} \\ \lvert refResult \rvert - 3.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	N/A
Both inputs are normalized or zero, intermediate result overflow, i.e. $\left\{ \begin{array}{c} (fltInY, fltInX) \in X \mid X \subset N \cup \{-0, +0\} \\ \frac{fltInY}{fltInX} \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, -Inf, Inf}

3.11 Function GFLIB_AtanYXShifted_FLT

3.11.1 Declaration

```
tFloat GFLIB_AtanYXShifted_FLT(tFloat fltInY, tFloat fltInX, const GFLIB_ATANYXSHIFTED_T_FLT
*pParam);
```

3.11.2 Arguments

Table 3-41. GFLIB_AtanYXShifted_FLT arguments

Type	Name	Direction	Description
tFloat	fltInY	**input**	The value of the first signal, assumed to be sin(θ).
tFloat	fltInX	**input**	The value of the second signal, assumed to be sin(θ + Δθ).
const GFLIB_ATANYXSHIFTED_T_FLT *	pParam	**input, output**	The parameters for the function.

3.11.3 Worst-Case Error Bounds

Let $(fltInX, fltInY, fltKx, fltKy, fltThetaAdj) \in X$ be a vector of inputs to GFLIB_AtanYXShifted_FLT,

$ir_1 = fltInY + fltInX$,

$ir_2 = ir_1 \cdot fltKy$,

$ir_3 = fltInX - fltInY$,

$ir_4 = ir_3 \cdot fltKx$,

$ir_6 = \mathrm{atan2}(ir_2, ir_4)$, where ir_2 is the y coordinate and ir_4 is the x coordinate and atan2(y, x) is the four-quadrant arctangent function,

$ir_7 = ir_6 - fltThetaAdj$,

$ie_1 = \max(\mathrm{fe}(ir_6 - 10 \cdot \mathrm{ULP}(ir_6)), \mathrm{fe}(ir_6 + 13.5 \cdot \mathrm{ULP}(ir_6)), \mathrm{fe}(fltThetaAdj))$,

$ie_2 = \mathrm{fe}(ir_7)$,

$$cb_1 = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me_1 = 0.5 + 13.5 \cdot 2^{cb_1}$,

$ie_3 = \max(\mathrm{fe}(ir_7 - me_1 \cdot \mathrm{ULP}(ir_7)), \mathrm{fe}(ir_7 + me_1 \cdot \mathrm{ULP}(ir_7)), \mathrm{fe}(2\pi))$,

$ie_4 = \mathrm{fe}(refResult)$,

$$cb_2 = \begin{cases} 0, & ie_3 - ie_4 \le 0 \\ ie_3 - ie_4, & ie_3 - ie_4 > 0 \end{cases},$$

$$me_2 = 0.5 + (me_1 + 0.5) \cdot 2^{cb_2},$$

$$me_3 = \begin{cases} me_2, & (ir_7 + me_1 \cdot \mathrm{ULP}(ir_7) > \pi) \text{ or } (ir_7 - me_1 \cdot \mathrm{ULP}(ir_7) < -\pi) \\ me_1, & otherwise \end{cases},$$

$refResult$ be the theoretical exact result,

then

Table 3-42. GFLIB_AtanYXShifted_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ (fltInX, fltInY, fltKx, fltKy, fltThetaAdj) \in X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2, ir_3, ir_4, ir_5, ir_6, ir_7\} \subset N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - me_3 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me_3 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -me_3, me_3 \rangle$	N/A
Input values which contain at least one NaN, Inf, or -Inf, i.e. $\left\{ (fltInX, fltInY, fltKx, fltKy, fltThetaAdj) \in X \;\middle|\; X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, -Inf, NaN)
Input values which contain at least one denormalized number while other inputs are normalized or zero, i.e. $\left\{ (fltInX, fltInY, fltKx, fltKy, fltThetaAdj) \in X \;\middle|\; X \cap M = \varnothing \wedge X \cap D \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, -Inf, NaN)
Normalized or zero input values which cause result underflow, i.e. $\left\{ (fltInX, fltInY, fltKx, fltKy, fltThetaAdj) \in X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me_3 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ (fltInX, fltInY, fltKx, fltKy, fltThetaAdj) \in X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \left\{ \begin{array}{c} ir_1, \lvert ir_2 \rvert + 1.5 \cdot \mathrm{ULP}(ir_2), \\ ir_3, \lvert ir_4 \rvert + 1.5 \cdot \mathrm{ULP}(ir_4), \\ \frac{\lvert ir_2 \rvert + 1.5 \cdot \mathrm{ULP}(ir_2)}{\lvert ir_4 \rvert - 1.5 \cdot \mathrm{ULP}(ir_4)} \end{array} \right\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}

Table continues on the next page...

Table 3-42. GFLIB_AtanYXShifted_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ (fltInX, fltInY, fltKx, fltKy, fltThetaAdj) \in X \;\middle|\; X \subset N \cup \{-0, +0\}, \left\{ ir_1, \lvert ir_2 \rvert - 1.5 \cdot \mathrm{ULP}(ir_2), ir_3, \lvert ir_4 \rvert - 1.5 \cdot \mathrm{ULP}(ir_4), \frac{\lvert ir_2 \rvert - 1.5 \cdot \mathrm{ULP}(ir_2)}{\lvert ir_4 \rvert + 1.5 \cdot \mathrm{ULP}(ir_4)}, \lvert ir_6 \rvert - 13.5 \cdot \mathrm{ULP}(ir_6), \lvert ir_7 \rvert - me_1 \cdot \mathrm{ULP}(ir_7), \right\} \cap D \neq \varnothing, ir_2 \neq 0, ir_4 \neq 0, ir_5 \neq 0, ir_6 \neq 0, ir_7 \neq 0 \right\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}

3.12 Function GFLIB_ControllerPIp_FLT

3.12.1 Declaration

```
tFloat GFLIB_ControllerPIp_FLT(tFloat fltInErr, GFLIB_CONTROLLER_PI_P_T_FLT *const pParam);
```

3.12.2 Arguments

Table 3-43. GFLIB_ControllerPIp_FLT arguments

Type	Name	Direction	Description
tFloat	fltInErr	**input**	Input error signal to the controller in single precision floating format.
GFLIB_CONTROLLER _PI_P_T_FLT *const	pParam	**input, output**	Pointer to the controller parameters structure.

3.12.3 Worst-Case Error Bounds

Let $(fltInErr, fltPropGain, fltIntegGain, fltIntegPartK\_1, fltInK\_1) \in X$ be a vector of inputs to GFLIB_ControllerPIp_FLT, *fltIntegPartK_1* input represents a real number with an error of max. ±0.5 ulp,

refResult be the theoretical exact result,

$ir_1 = fltInErr + fltInK\_1$,

$ir_2 = ir_1 \cdot fltIntegGain$,

$ie_1 = \max(\mathrm{fe}(|ir_2| + \mathrm{ULP}(ir_2)), \mathrm{fe}(|fltIntegPartK\_1| + 0.5 \cdot \mathrm{ULP}(fltIntegPartK\_1)))$,

$refInteg = fltIntegPartK\_1 + ir_2$,

$ie_2 = \mathrm{fe}(refInteg)$,

$$cb_1 = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me_1 = 0.5 + 1.5 \cdot 2^{cb_1}$,

$ir_3 = fltInErr \cdot fltPropGain$,

$ie_3 = \max(\mathrm{fe}(ir_3), \mathrm{fe}(|refInteg| + me_1 \cdot \mathrm{ULP}(refInteg)))$,

$ie_4 = \mathrm{fe}(refResult)$,

$$cb_2 = \begin{cases} 0, & ie_3 - ie_4 \le 0 \\ ie_3 - ie_4, & ie_3 - ie_4 > 0 \end{cases},$$

$me_2 = 0.5 + me_1 \cdot 2^{cb_2}$,

then

Table 3-44. GFLIB_ControllerPIp_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \;\middle|\; X \cap D \ne \varnothing \wedge X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \;\middle|\; X \cap M \ne \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e.	$\langle -me_2, me_2 \rangle$	N/A

Table continues on the next page...

Table 3-44. GFLIB_ControllerPIp_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \middle\| \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2, ir_3\} \subset N \cup \{-0, +0\}, \\ refInteg \in N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \begin{pmatrix} \|refResult\| - me_2 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \|refResult\| + me_2 \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{array} \right\}$		
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \middle\| \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \|refResult\| - me_2 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \middle\| \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \|refResult\| + me_2 \cdot \mathrm{ULP}(refResult) = Inf \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \middle\| \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \left\{ \begin{array}{c} \|ir_1\| - 0.5 \cdot \mathrm{ULP}(ir_1), \\ \|ir_2\| - \mathrm{ULP}(ir_2), \\ \|refInteg\| - me_1 \cdot \mathrm{ULP}(refInteg), \\ ir_3 \end{array} \right\} \cap D \neq \emptyset \wedge \\ \{ir_1, ir_2, ir_3, refInteg\} \notin \{-0, +0\} \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result overflow, i.e.	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-44. GFLIB_ControllerPIp_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array}\right) \in X \;\middle\vert\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \left\{\begin{array}{c} \lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1), \\ \lvert ir_2 \rvert + \mathrm{ULP}(ir_2), \\ \lvert refInteg \rvert + me_1 \cdot \mathrm{ULP}(refInteg), \\ ir_3 \end{array}\right\} \cap M \neq \emptyset \end{array} \right\}$		

Table 3-45. GFLIB_ControllerPIp_FLT Worst-Case Error Bounds - fltIntegPartK_1 Output

Subset of input domain	Worst-case error bounds for fltIntegPartK_1 output [ulp]	Allowed specific values (regardless the error bounds) for fltIntegPartK_1 output
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array}\right) \in X \;\middle\vert\; X \cap D \neq \emptyset \wedge X \cap M = \emptyset \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array}\right) \in X \;\middle\vert\; X \cap M \neq \emptyset \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array}\right) \in X \;\middle\vert\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2\} \subset N \cup \{-0, +0\}, \\ refInteg = 0 \vee \\ \left(\begin{array}{c} \lvert refInteg \rvert - me_1 \cdot \mathrm{ULP}(refInteg) \notin D \wedge \\ \lvert refInteg \rvert + me_1 \cdot \mathrm{ULP}(refInteg) \neq Inf \end{array}\right) \end{array} \right\}$	$\langle -me_1, me_1 \rangle$	N/A
Normalized or zero input values which cause *fltIntegPartK_1* output underflow, i.e.	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-45. GFLIB_ControllerPIp_FLT Worst-Case Error Bounds - fltIntegPartK_1 Output (continued)

Subset of input domain	Worst-case error bounds for fltIntegPartK_1 output [ulp]	Allowed specific values (regardless the error bounds) for fltIntegPartK_1 output
$\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \middle\| \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \|refInteg\| - me_1 \cdot \mathrm{ULP}(refInteg) \in D \wedge \\ refInteg \neq 0 \end{array} \right\}$		
Normalized or zero input values which cause *fltIntegPartK_1* output overflow, i.e. $\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \middle\| \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \|refInteg\| + me_1 \cdot \mathrm{ULP}(refInteg) = Inf \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \middle\| \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} \|ir_1\| - 0.5 \cdot \mathrm{ULP}(ir_1), \\ \|ir_2\| - \mathrm{ULP}(ir_2), \end{array} \right\} \cap D \neq \varnothing \wedge \\ \{ir_1, ir_2\} \notin \{-0,\ +0\} \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \middle\| \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} \|ir_1\| + 0.5 \cdot \mathrm{ULP}(ir_1), \\ \|ir_2\| + \mathrm{ULP}(ir_2), \end{array} \right\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-46. GFLIB_ControllerPIp_FLT Worst-Case Error Bounds - fltInK_1 Output

Subset of input domain	Worst-case error bounds for fltInK_1 output [ulp]	Allowed specific values (regardless the error bounds) for fltInK_1 output
Entire input domain	0	*fltInErr*

3.13 Function GFLIB_ControllerPIpAW_FLT

3.13.1 Declaration

```
tFloat GFLIB_ControllerPIpAW_FLT(tFloat fltInErr, GFLIB_CONTROLLER_PIAW_P_T_FLT *const
pParam);
```

3.13.2 Arguments

Table 3-47. GFLIB_ControllerPIpAW_FLT arguments

Type	Name	Direction	Description
tFloat	fltInErr	**input**	Input error signal to the controller is a single precision floating point data type.
GFLIB_CONTROLLER_PIAW_P_T_FLT *const	pParam	**input, output**	Pointer to the controller parameters structure.

3.13.3 Worst-Case Error Bounds

Let (*fltInErr*, *fltPropGain*, *fltIntegGain*, *fltLowerLimit*, *fltUpperLimit*, *fltIntegPartK_1*, *fltInK_1*) $\in X$ be a vector of inputs to GFLIB_ControllerPIpAW_FLT,

fltIntegPartK_1 input represents a real number with an error of max. ±0.5 ulp,

refResult be the theoretical exact result,

$ir_1 = fltInErr + fltInK\_1$,

$ir_2 = ir_1 \cdot fltIntegGain$,

$ie_1 = \max(\mathrm{fe}(|ir_2| + \mathrm{ULP}(ir_2)), \mathrm{fe}(|fltIntegPartK\_1| + 0.5 \cdot \mathrm{ULP}(fltIntegPartK\_1)))$,

$refInteg = fltIntegPartK\_1 + ir_2$,

$ie_2 = \mathrm{fe}(refInteg)$,

$$cb_1 = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$$me_1 = \begin{cases} 0, & refInteg + \left(0.5 + 1.5 \cdot 2^{cb_1}\right) \cdot \mathrm{ULP}(refInteg) < fltLowerLimit \\ 0, & refInteg - \left(0.5 + 1.5 \cdot 2^{cb_1}\right) \cdot \mathrm{ULP}(refInteg) > fltUpperLimit \\ 0.5 + 1.5 \cdot 2^{cb_1}, & \text{otherwise} \end{cases},$$

$$ir_3 = fltInErr \cdot fltPropGain,$$

$$ie_3 = \max(\mathrm{fe}(ir_3), \mathrm{fe}(|refInteg| + me_1 \cdot \mathrm{ULP}(refInteg))),$$

$$ie_4 = \mathrm{fe}(refResult),$$

$$cb_2 = \begin{cases} 0, & ie_3 - ie_4 \le 0 \\ ie_3 - ie_4, & ie_3 - ie_4 > 0 \end{cases},$$

$$me_2 = 0.5 + me_1 \cdot 2^{cb_2},$$

then

Table 3-48. GFLIB_ControllerPIpAW_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltLowerLimit, fltUpperLimit, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \;\middle|\; X \cap D \neq \varnothing \wedge X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltLowerLimit, fltUpperLimit, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \;\middle|\; X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e.	$\langle -me_2, me_2 \rangle$	N/A

Table continues on the next page...

Table 3-48. GFLIB_ControllerPIpAW_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \left(\begin{matrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{matrix} \right) \in X \middle\vert \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2, ir_3\} \subset N \cup \{-0,\ +0\}, \\ refInteg \in N \cup \{-0,\ +0\}, \\ refResult = 0 \vee \\ \left(\begin{matrix} \lvert refResult \rvert - me_2 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me_2 \cdot \mathrm{ULP}(refResult) \neq Inf \end{matrix} \right) \end{matrix} \right\}$		
Normalized or zero input values which cause result underflow, i.e. $\left\{ \left(\begin{matrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{matrix} \right) \in X \middle\vert \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me_2 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause result overflow, i.e. $\left\{ \left(\begin{matrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{matrix} \right) \in X \middle\vert \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert + me_2 \cdot \mathrm{ULP}(refResult) = Inf \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e.	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table continues on the next page...

Table 3-48. GFLIB_ControllerPIpAW_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{pmatrix} \in X \,\middle\vert\, \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \left\{ \begin{array}{c} \lvert ir_1 \rvert - 0.5 \cdot \mathrm{ULP}(ir_1), \\ \lvert ir_2 \rvert - \mathrm{ULP}(ir_2), \\ \lvert refInteg \rvert - me_1 \cdot \mathrm{ULP}(refInteg), \\ ir_3 \end{array} \right\} \cap D \neq \varnothing \land \\ \{ir_1, ir_2, ir_3, refInteg\} \notin \{-0, +0\} \end{array} \right\}$		
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{pmatrix} \in X \,\middle\vert\, \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \left\{ \begin{array}{c} \lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1), \\ \lvert ir_2 \rvert + \mathrm{ULP}(ir_2), \\ \lvert refInteg \rvert + me_1 \cdot \mathrm{ULP}(refInteg), \\ ir_3 \end{array} \right\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-49. GFLIB_ControllerPIpAW_FLT Worst-Case Error Bounds - fltIntegPartK_1 Output

Subset of input domain	Worst-case error bounds for fltIntegPartK_1 output [ulp]	Allowed specific values (regardless the error bounds) for fltIntegPartK_1 output
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{pmatrix} \in X \,\middle\vert\, X \cap D \neq \varnothing \land X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e.	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

Table continues on the next page...

Table 3-49. GFLIB_ControllerPIpAW_FLT Worst-Case Error Bounds - fltIntegPartK_1 Output (continued)

Subset of input domain	Worst-case error bounds for fltIntegPartK_1 output [ulp]	Allowed specific values (regardless the error bounds) for fltIntegPartK_1 output
$\left\{ \left. \begin{pmatrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{pmatrix} \in X \right\| \begin{matrix} \\ X \cap M \neq \emptyset \end{matrix} \right\}$		
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \left. \begin{pmatrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{pmatrix} \in X \right\| \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1,\ ir_2\} \subset N \cup \{-0,\ +0\}, \\ refInteg = 0 \vee \\ \begin{pmatrix} \lvert refInteg \rvert - me_1 \cdot \mathrm{ULP}(refInteg) \notin D \wedge \\ \lvert refInteg \rvert + me_1 \cdot \mathrm{ULP}(refInteg) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -me_1,\ me_1 \rangle$	N/A
Normalized or zero input values which cause *fltIntegPartK_1* output underflow, i.e. $\left\{ \left. \begin{pmatrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{pmatrix} \in X \right\| \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \lvert refInteg \rvert - me_1 \cdot \mathrm{ULP}(refInteg) \in D \wedge \\ refInteg \neq 0 \end{matrix} \right\}$	$(-Inf,\ Inf)$	N/A
Normalized or zero input values which cause *fltIntegPartK_1* output overflow, i.e. $\left\{ \left. \begin{pmatrix} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{pmatrix} \in X \right\| \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \lvert refInteg \rvert + me_1 \cdot \mathrm{ULP}(refInteg) = Inf \end{matrix} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e.	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}

Table continues on the next page...

Table 3-49. GFLIB_ControllerPIpAW_FLT Worst-Case Error Bounds - fltIntegPartK_1 Output (continued)

Subset of input domain	Worst-case error bounds for fltIntegPartK_1 output [ulp]	Allowed specific values (regardless the error bounds) for fltIntegPartK_1 output
$\left\{ \begin{array}{c} \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} \lvert ir_1 \rvert - 0.5 \cdot \mathrm{ULP}(ir_1), \\ \lvert ir_2 \rvert - \mathrm{ULP}(ir_2), \end{array} \right\} \cap D \neq \varnothing \wedge \\ \{ir_1,\ ir_2\} \notin \{-0,\ +0\} \end{array} \right\}$		
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \begin{array}{c} \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltLowerLimit,\ fltUpperLimit, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} \lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1), \\ \lvert ir_2 \rvert + \mathrm{ULP}(ir_2), \end{array} \right\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}

Table 3-50. GFLIB_ControllerPIpAW_FLT Worst-Case Error Bounds - fltInK_1 Output

Subset of input domain	Worst-case error bounds for fltInK_1 output [ulp]	Allowed specific values (regardless the error bounds) for fltInK_1 output
Entire input domain	0	*fltInErr*

3.14 Function GFLIB_ControllerPIr_FLT

3.14.1 Declaration

```
tFloat GFLIB_ControllerPIr_FLT(tFloat fltInErr, GFLIB_CONTROLLER_PI_R_T_FLT *const pParam);
```

3.14.2 Arguments

Table 3-51. GFLIB_ControllerPIr_FLT arguments

Type	Name	Direction	Description
tFloat	fltInErr	**input**	Input error signal to the controller as a single precision floating point value.
GFLIB_CONTROLLER_PI_R_T_FLT *const	pParam	**input, output**	Pointer to the controller parameters structure.

3.14.3 Worst-Case Error Bounds

Let $(fltInErr, fltCC1sc, fltCC2sc, fltAcc, fltInErrK1) \in X$ be a vector of inputs to GFLIB_ControllerPIr_FLT,

$fltAcc$ input represents a real number with an error of max. ±0.5 ulp,

$refResult$ be the theoretical exact result,

$ir_1 = fltInErr \cdot fltCC1sc$,

$ir_2 = fltInErrK1 \cdot fltCC2sc$,

$ir_3 = ir_1 + ir_2$,

$ie_1 = \max(\mathrm{fe}(|ir_3| + \mathrm{ULP}(ir_3)), \mathrm{fe}(|fltAcc| + 0.5 \cdot \mathrm{ULP}(fltAcc)))$,

$ie_2 = \mathrm{fe}(refResult)$,

$$cb = \begin{cases} 0, & ie_1 - ie_2 \leq 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me = 0.5 + 1.5 \cdot 2^{cb}$,

then

Table 3-52. GFLIB_ControllerPIr_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \left(\begin{matrix} fltInErr, fltCC1sc, fltCC2sc, \\ fltAcc, fltInErrK1 \end{matrix} \right) \in X \;\middle|\; X \cap D \neq \varnothing \wedge X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-52. GFLIB_ControllerPIr_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \;\middle|\; X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2, ir_3\} \subset N \cup \{-0,\ +0\}, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -me,\ me \rangle$	N/A
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf,\ Inf)$	N/A
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) = Inf \end{matrix} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2\} \cap D \neq \varnothing \vee \\ (\lvert ir_3 \rvert - \mathrm{ULP}(ir_3) \in D \wedge ir_3 \neq 0) \end{matrix} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result overflow, i.e.	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}

Table 3-52. GFLIB_ControllerPIr_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \middle\| \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2, \lvert ir_3\rvert + \mathrm{ULP}(ir_3)\} \cap M \neq \varnothing \end{matrix} \right\}$		

Table 3-53. GFLIB_ControllerPIr_FLT Worst-Case Error Bounds - fltAcc Output

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \middle\| X \cap D \neq \varnothing \wedge X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \middle\| X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \middle\| \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2, ir_3\} \subset N \cup \{-0,\ +0\}, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult\rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult\rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause *fltAcc* output underflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAcc,\ fltInErrK1 \end{pmatrix} \in X \middle\| \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult\rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-53. GFLIB_ControllerPIr_FLT Worst-Case Error Bounds - fltAcc Output (continued)

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
Normalized or zero input values which cause *fltAcc* output overflow, i.e. $\left\{ \begin{pmatrix} fltInErr, fltCC1sc, fltCC2sc, \\ fltAcc, fltInErrK1 \end{pmatrix} \in X \middle\vert \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) = Inf \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{pmatrix} fltInErr, fltCC1sc, fltCC2sc, \\ fltAcc, fltInErrK1 \end{pmatrix} \in X \middle\vert \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2\} \cap D \neq \varnothing \vee \\ (\lvert ir_3 \rvert - \mathrm{ULP}(ir_3) \in D \wedge ir_3 \neq 0) \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \begin{pmatrix} fltInErr, fltCC1sc, fltCC2sc, \\ fltAcc, fltInErrK1 \end{pmatrix} \in X \middle\vert \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2, \lvert ir_3 \rvert + \mathrm{ULP}(ir_3)\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-54. GFLIB_ControllerPIr_FLT Worst-Case Error Bounds - fltInErrK1 Output

Subset of input domain	Worst-case error bounds for fltInErrK1 output [ulp]	Allowed specific values (regardless the error bounds) for fltInErrK1 output
Entire input domain	0	*fltInErr*

3.15 Function GFLIB_ControllerPIrAW_FLT

3.15.1 Declaration

```
tFloat GFLIB_ControllerPIrAW_FLT(tFloat fltInErr, GFLIB_CONTROLLER_PIAW_R_T_FLT *const
pParam);
```

3.15.2 Arguments

Table 3-55. GFLIB_ControllerPIrAW_FLT arguments

Type	Name	Direction	Description
tFloat	fltInErr	**input**	Input error signal to the controller in single precision floating point data format.
GFLIB_CONTROLLER_PIAW_R_T_FLT *const	pParam	**input, output**	Pointer to the controller parameters structure.

3.15.3 Worst-Case Error Bounds

Let $(fltInErr, fltCC1sc, fltCC2sc, fltAcc, fltInErrK1, fltUpperLimit, fltLowerLimit) \in X$ be a vector of inputs to GFLIB_ControllerPIrAW_FLT,

$fltAcc$ input represents a real number with an error of max. ±0.5 ulp,

$refResult$ be the theoretical exact result,

$ir_1 = fltInErr \cdot fltCC1s$,

$ir_2 = fltInErrK1 \cdot fltCC2sc$,

$ir_3 = ir_1 + ir_2$,

$ie_1 = \max(\mathrm{fe}(|ir_3| + \mathrm{ULP}(ir_3)), \mathrm{fe}(|fltAcc| + 0.5 \cdot \mathrm{ULP}(fltAcc)))$,

$ie_2 = \mathrm{fe}(refResult)$,

$$cb = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me = 0.5 + 1.5 \cdot 2^{cb}$,

then

Table 3-56. GFLIB_ControllerPIrAW_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\| \begin{matrix} \\ X \cap D \neq \varnothing \wedge X \cap M = \varnothing \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\| \begin{matrix} \\ X \cap M \neq \varnothing \end{matrix} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\| \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2, ir_3\} \subset N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\| \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\| \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) = Inf \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e.	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table continues on the next page...

Table 3-56. GFLIB_ControllerPIrAW_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\vert \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2\} \cap D \neq \varnothing \lor \\ (\lvert ir_3 \rvert - \mathrm{ULP}(ir_3) \in D \land ir_3 \neq 0) \end{array} \right\}$		
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\vert \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2, \lvert ir_3 \rvert + \mathrm{ULP}(ir_3)\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-57. GFLIB_ControllerPIrAW_FLT Worst-Case Error Bounds - fltAcc Output

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\vert X \cap D \neq \varnothing \land X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \middle\vert X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e.	$\langle -me, me \rangle$	N/A

Table continues on the next page...

Table 3-57. GFLIB_ControllerPIrAW_FLT Worst-Case Error Bounds - fltAcc Output (continued)

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
$\left\{ \begin{array}{c} \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2, ir_3\} \subset N \cup \{-0,\ +0\}, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$		
Normalized or zero input values which cause *fltAcc* output underflow, i.e. $\left\{ \begin{array}{c} \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause *fltAcc* output overflow, i.e. $\left\{ \begin{array}{c} \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) = Inf \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{array}{c} \begin{pmatrix} fltInErr,\ fltCC1sc,\ fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit,\ fltLowerLimit \end{pmatrix} \in X \\ X \subset N \cup \{-0,\ +0\}, \\ \{ir_1, ir_2\} \cap D \neq \varnothing \vee \\ (\lvert ir_3 \rvert - \mathrm{ULP}(ir_3) \in D \wedge ir_3 \neq 0) \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result overflow, i.e.	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-57. GFLIB_ControllerPIrAW_FLT Worst-Case Error Bounds - fltAcc Output

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
$\left\{ \left(\begin{array}{c} fltInErr, fltCC1sc, fltCC2sc, \\ fltAccfltInErrK1, \\ fltUpperLimit, fltLowerLimit \end{array} \right) \in X \middle\| \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2, \|ir_3\| + \mathrm{ULP}(ir_3)\} \cap M \neq \varnothing \end{array} \right\}$		

Table 3-58. GFLIB_ControllerPIrAW_FLT Worst-Case Error Bounds - fltInErrK1 Output

Subset of input domain	Worst-case error bounds for fltAcc output [ulp]	Allowed specific values (regardless the error bounds) for fltAcc output
Entire input domain	0	*fltInErr*

3.16 Function GFLIB_Cos_FLT

3.16.1 Declaration

```
tFloat GFLIB_Cos_FLT(tFloat fltIn, const GFLIB_COS_T_FLT *const pParam);
```

3.16.2 Arguments

Table 3-59. GFLIB_Cos_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument is a single precision floating point number that contains an angle in radians between [- π, π].
const GFLIB_COS_T_FLT *const	pParam	**input**	Pointer to an array of approximation coefficients.

3.16.3 Worst-Case Error Bounds

Only the default approximation coefficients are considered.

Let $fltIn \in X$ be an input to GFLIB_Cos_FLT,

then

Table 3-60. GFLIB_Cos_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input NaN or infinity, i.e. $\{fltIn \in X \mid fltIn \in M\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Input denormalized, normalized or zero, and within the allowed input interval, i.e. $\left\{ fltIn \in X \;\middle|\; \begin{array}{l} X \cap M = \emptyset, \\ -\pi \le fltIn \le \pi \end{array} \right\}$	$\langle -334, 334 \rangle$	N/A
Input normalized and beyond the allowed range, i.e. $\left\{ fltIn \in X \;\middle|\; \begin{array}{l} X \cap M = \emptyset, \\ fltIn < -\pi \vee fltIn > \pi \end{array} \right\}$	$(-Inf, Inf)$	{NaN, -Inf, Inf}

3.17 Function GFLIB_Hyst_FLT

3.17.1 Declaration

```
tFloat GFLIB_Hyst_FLT(tFloat fltIn, GFLIB_HYST_T_FLT *const pParam);
```

3.17.2 Arguments

Table 3-61. GFLIB_Hyst_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input value, in single precision floating point data format.
GFLIB_HYST_T_FLT *const	pParam	**input, output**	Pointer to the structure with parameters and states of the hysteresis function. Arguments of the structure contain a single precision floating point values.

3.17.3 Worst-Case Error Bounds

Let $(fltIn, fltHystOn, fltHystOff, fltOutValOn, fltOutValOff) \in X$ be a vector of inputs to GFLIB_Hyst_FLT,

then

Table 3-62. GFLIB_Hyst_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs are normalized, denormalized, or zero, and *fltHystOn* is greater than *fltHystOff*, i.e. $\left\{ \left(\begin{matrix} fltIn, fltHystOn, fltHystOff, \\ fltOutValOn, fltOutValOff \end{matrix} \right) \in X \;\middle\|\; \begin{matrix} X \cap M = \emptyset \wedge \\ fltHystOff < fltHystOn \end{matrix} \right\}$	0	N/A
At least one of the inputs is NaN or infinite, or *fltHystOn* is not greater than *fltHystOff*, i.e. $\left\{ \left(\begin{matrix} fltIn, fltHystOn, fltHystOff, \\ fltOutValOn, fltOutValOff \end{matrix} \right) \in X \;\middle\|\; \begin{matrix} X \cap M \neq \emptyset \vee \\ fltHystOff \geq fltHystOn \end{matrix} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

3.18 Function GFLIB_IntegratorTR_FLT

3.18.1 Declaration

```
tFloat GFLIB_IntegratorTR_FLT(tFloat fltIn, GFLIB_INTEGRATOR_TR_T_FLT *const pParam);
```

3.18.2 Arguments

Table 3-63. GFLIB_IntegratorTR_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument to be integrated.
GFLIB_INTEGRATOR_TR_T_FLT *const	pParam	**input, output**	Pointer to the integrator parameters structure.

3.18.3 Worst-Case Error Bounds

Let $(fltIn, fltState, fltInK1, fltC1) \in X$ be a vector of inputs to GFLIB_IntegratorTR_FLT,

fltState input represents a real number with an error of max. ±0.5 ulp,

refResult be the theoretical exact result,

$ir_1 = fltIn + fltInK1$,

$ir_2 = ir_1 \cdot fltC1$,

$ie_1 = \max(\mathrm{fe}(|ir_2| + \mathrm{ULP}(ir_2)), \mathrm{fe}(|fltState| + 0.5 \cdot \mathrm{ULP}(fltState)))$,

$ie_2 = \mathrm{fe}(refResult)$,

$$cb = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me = 0.5 + 1.5 \cdot 2^{cb}$,

then

Table 3-64. GFLIB_IntegratorTR_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{array}{c} (fltIn, fltState, fltInK1, fltC1) \in X \mid \\ X \cap D \ne \varnothing \wedge X \cap M = \varnothing \end{array} \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{array}{c} (fltIn, fltState, fltInK1, fltC1) \in X \mid \\ X \cap M \ne \varnothing \end{array} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{array}{c} (fltIn, fltState, fltInK1, fltC1) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2\} \subset N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \left(\begin{array}{c} |refResult| - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ |refResult| + me \cdot \mathrm{ULP}(refResult) \ne Inf \end{array} \right) \end{array} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause result underflow, i.e.		N/A

Table continues on the next page...

Table 3-64. GFLIB_IntegratorTR_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	
Normalized or zero input values which cause result overflow, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) = Inf \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} (\lvert ir_1 \rvert - 0.5 \cdot \mathrm{ULP}(ir_1)), \\ (\lvert ir_2 \rvert - \mathrm{ULP}(ir_2)) \end{array} \right\} \cap D \neq \varnothing \vee \\ ir_1 \neq 0 \vee ir_2 \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} (\lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1)), \\ (\lvert ir_2 \rvert + \mathrm{ULP}(ir_2)) \end{array} \right\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-65. GFLIB_IntegratorTR_FLT Worst-Case Error Bounds - fltState Output

Subset of input domain	Worst-case error bounds for fltState output [ulp]	Allowed specific values (regardless the error bounds) for fltState output
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \cap D \neq \varnothing \wedge X \cap M = \varnothing \end{array} \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \cap M \neq \varnothing \end{array} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

Table continues on the next page...

Table 3-65. GFLIB_IntegratorTR_FLT Worst-Case Error Bounds - fltState Output (continued)

Subset of input domain	Worst-case error bounds for fltState output [ulp]	Allowed specific values (regardless the error bounds) for fltState output
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \{ir_1,\ ir_2\} \subset N \cup \{-0,\ +0\}, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -me,\ me \rangle$	N/A
Normalized or zero input values which cause *fltState* output underflow, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf,\ Inf)$	N/A
Normalized or zero input values which cause *fltState* output overflow, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) = Inf \end{array} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} (\lvert ir_1 \rvert - 0.5 \cdot \mathrm{ULP}(ir_1)), \\ (\lvert ir_2 \rvert - \mathrm{ULP}(ir_2)) \end{array} \right\} \cap D \neq \varnothing \vee \\ ir_1 \neq 0 \vee ir_2 \neq 0 \end{array} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate result overflow, i.e. $\left\{ \begin{array}{c} (fltIn,\ fltState,\ fltInK1,\ fltC1) \in X \mid \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} (\lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1)), \\ (\lvert ir_2 \rvert + \mathrm{ULP}(ir_2)) \end{array} \right\} \cap M \neq \varnothing \end{array} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}

Table 3-66. GFLIB_IntegratorTR_FLT Worst-Case Error Bounds - fltInK1 Output

Subset of input domain	Worst-case error bounds for fltInK1 output [ulp]	Allowed specific values (regardless the error bounds) for fltInK1 output
Entire input domain	0	*fltIn*

3.19 Function GFLIB_Limit_FLT

3.19.1 Declaration

```
tFloat GFLIB_Limit_FLT(tFloat fltIn, const GFLIB_LIMIT_T_FLT *const pParam);
```

3.19.2 Arguments

Table 3-67. GFLIB_Limit_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input value.
const GFLIB_LIMIT_T_FLT *const	pParam	**input**	Pointer to the limits structure.

3.19.3 Worst-Case Error Bounds

Let $(fltIn, fltLowerLimit, fltUpperLimit) \in X$ be a vector of inputs to GFLIB_Limit_FLT, then

Table 3-68. GFLIB_Limit_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs are normalized, denormalized, or zero, and *fltUpperLimit* is greater than *fltLowerLimit*, i.e.	0	N/A

Table continues on the next page...

Table 3-68. GFLIB_Limit_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ (fltIn, fltLowerLimit, fltUpperLimit) \in X \;\middle|\; \begin{array}{c} X \cap M = \varnothing \wedge \\ fltLowerLimit < fltUpperLimit \end{array} \right\}$		
At least one of the inputs is NaN or infinite, or *fltUpperLimit* is not greater than *fltLowerLimit*, i.e. $\left\{ (fltIn, fltLowerLimit, fltUpperLimit) \in X \;\middle|\; \begin{array}{c} X \cap M \neq \varnothing \vee \\ fltLowerLimit \geq fltUpperLimit \end{array} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

3.20 Function GFLIB_LowerLimit_FLT

3.20.1 Declaration

```
tFloat GFLIB_LowerLimit_FLT(tFloat fltIn, const GFLIB_LOWERLIMIT_T_FLT *const pParam);
```

3.20.2 Arguments

Table 3-69. GFLIB_LowerLimit_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input value.
const GFLIB_LOWERLIMIT_T_FLT *const	pParam	**input**	Pointer to the limits structure.

3.20.3 Worst-Case Error Bounds

Let $(fltIn, fltLowerLimit) \in X$ be a vector of inputs to GFLIB_LowerLimit_FLT, then

Table 3-70. GFLIB_LowerLimit_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs are normalized, denormalized, or zero, i.e. $\left\{ (fltIn, fltLowerLimit) \in X \;\middle|\; X \cap M = \varnothing \right\}$	0	N/A
At least one of the inputs is NaN or infinite, i.e. $\left\{ (fltIn, fltLowerLimit) \in X \;\middle|\; X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

3.21 Function GFLIB_Lut1D_FLT

3.21.1 Declaration

```
tFloat GFLIB_Lut1D_FLT(tFloat fltIn, const GFLIB_LUT1D_T_FLT *const pParam);
```

3.21.2 Arguments

Table 3-71. GFLIB_Lut1D_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	The abscissa for which 1D interpolation is performed.
const GFLIB_LUT1D_T_FLT *const	pParam	**input**	Pointer to the parameters structure with parameters of the look-up table function.

3.21.3 Worst-Case Error Bounds

Let $(fltIn, u32ShamOffset, \boldsymbol{fltTable}) \in X$ be a vector of inputs to GFLIB_Lut1D_FLT, $refResult$ be the theoretical exact result,

$$k = \begin{cases} fltIn \cdot 2^{u32ShamOffset} - \mathrm{fix}\left(fltIn \cdot 2^{u32ShamOffset}\right), & fltIn \geq 0 \\ fltIn \cdot 2^{u32ShamOffset} - \mathrm{fix}\left(fltIn \cdot 2^{u32ShamOffset}\right) + 1, & fltIn < 0 \end{cases},$$

$$mr_n = k \cdot \left(fltTable_{n+1} - fltTable_n\right),$$

$$ie = \max_n\left(\max\left(\mathrm{fe}\left(|mr_n| + 2.5 \cdot \mathrm{ULP}(mr_n) \right), \mathrm{fe}\left(fltTable_n \right) \right) - \mathrm{fe}\left(mr_n + fltTable_n \right) \right),$$

$$cb = \begin{cases} 0, & ie \le 0 \\ ie, & ie > 0 \end{cases},$$

$$me = 0.5 + 2.5 \cdot 2^{cb},$$

then

Table 3-72. GFLIB_Lut1D_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ (fltIn, u32ShamOffset, \boldsymbol{fltTable}) \in X \;\middle|\; X \cap D \neq \varnothing \wedge X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, or the input is beyond allowed range, i.e. $\left\{ (fltIn, u32ShamOffset, \boldsymbol{fltTable}) \in X \;\middle|\; X \cap M \neq \varnothing \vee \lvert fltIn \rvert \ge 1 \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, input within allowed range, i.e. $\left\{ \begin{array}{c} (fltIn, u32ShamOffset, \boldsymbol{fltTable}) \in X \;\middle|\; \\ X \subset N \cup \{-0, +0\}, \\ fltTable \subset N \cup \{-0, +0\} \\ \lvert fltIn \rvert < 1, \\ \max_n(\lvert mr_n \rvert) + 2.5 \cdot \mathrm{ULP}\left(\max_n(mr_n)\right) \neq Inf, \\ \bigvee mr_n = 0 \vee \lvert mr_n \rvert - 2.5 \cdot \mathrm{ULP}(mr_n) \notin D, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - me_2 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me_2 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{array}{c} (fltIn, u32ShamOffset, \boldsymbol{fltTable}) \in X \;\middle|\; \\ X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause intermediate result underflow, i.e.	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table 3-72. GFLIB_Lut1D_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} (fltIn, u32ShamOffset, \mathbf{fltTable}) \in X \;\middle\|\; \\ X \subset N \cup \{-0, +0\}, \\ \exists \lvert mr_n \rvert - 2.5 \cdot \mathrm{ULP}(mr_n) \in D \wedge mr_n \neq 0 \end{array} \right\}$		

3.22 Function GFLIB_Lut2D_FLT

3.22.1 Declaration

```
tFloat GFLIB_Lut2D_FLT(tFloat fltIn1, tFloat fltIn2, const GFLIB_LUT2D_T_FLT *const pParam);
```

3.22.2 Arguments

Table 3-73. GFLIB_Lut2D_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn1	**input**	First input variable for which 2D interpolation is performed. Input value is in single precision floating data format.
tFloat	fltIn2	**input**	Second input variable for which 2D interpolation is performed. Input value is in single precision floating data format.
const GFLIB_LUT2D_T_FLT *const	pParam	**input**	Pointer to the parameters structure with parameters of the two dimensional look-up table function.

3.22.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2, u32ShamOffset1, u32ShamOffset2, \mathbf{fltTable}) \in X$ be a vector of inputs to GFLIB_Lut2D_FLT,

refResult be the theoretical exact result,

$$kx = \begin{cases} fltIn1 \cdot 2^{u32ShamOffset1} - \mathrm{fix}\left(fltIn1 \cdot 2^{u32ShamOffset1}\right), & fltIn1 \geq 0 \\ fltIn1 \cdot 2^{u32ShamOffset1} - \mathrm{fix}\left(fltIn1 \cdot 2^{u32ShamOffset1}\right) + 1, & fltIn1 < 0 \end{cases},$$

$$ky = \begin{cases} fltIn2 \cdot 2^{u32ShamOffset2} - \mathrm{fix}\left(fltIn2 \cdot 2^{u32ShamOffset2}\right), & fltIn2 \geq 0 \\ fltIn2 \cdot 2^{u32ShamOffset2} - \mathrm{fix}\left(fltIn2 \cdot 2^{u32ShamOffset2}\right) + 1, & fltIn2 < 0 \end{cases},$$

$$yr = 2^{u32ShamOffset2} + 1,$$

$$z11_n = fltTable_n,$$

$$z21_n = fltTable_{n+yr},$$

$$z12_n = fltTable_{n+1},$$

$$z22_n = fltTable_{n+1+yr},$$

$$mra_n = kx \cdot (z21_n - z11_n),$$

$$sra = \max_n\left(\max\left(\mathrm{fe}\left(|mra_n| + 2.5 \cdot \mathrm{ULP}(mra_n)\right), \mathrm{fe}(z11_n)\right) - \mathrm{fe}\left(mra_n + z11_n\right)\right),$$

$$cb_1 = \begin{cases} 0, & sra \leq 0 \\ sra, & sra > 0 \end{cases},$$

$$me_1 = 0.5 + 2.5 \cdot 2^{cb_1},$$

$$mrb_n = kx \cdot (z22_n - z12_n),$$

$$srb = \max_n\left(\max\left(\mathrm{fe}\left(|mrb_n| + 2.5 \cdot \mathrm{ULP}(mrb_n)\right), \mathrm{fe}(z12_n)\right) - \mathrm{fe}\left(mrb_n + z12_n\right)\right),$$

$$cb_2 = \begin{cases} 0, & srb \leq 0 \\ srb, & srb > 0 \end{cases},$$

$$me_2 = 0.5 + 2.5 \cdot 2^{cb_2},$$

$$mrc_n = ky \cdot \left(\left(mrb_n + z12_n\right) - \left(mra_n + z11_n\right)\right),$$

$$mec = \left(2 \cdot (me_1 + me_2 + 0.5) + 0.5\right) \cdot \mathrm{ULP}(mrc_n),$$

$$src = \max_n\left(\max\left(\mathrm{fe}\left(|mrc_n| + mec\right), \mathrm{fe}\left(mra_n + z11_n\right)\right) - \mathrm{fe}\left(mrc_n + \left(mra_n + z11_n\right)\right)\right),$$

$$cb_3 = \begin{cases} 0, & src \leq 0 \\ src, & src > 0 \end{cases},$$

$$me_3 = 0.5 + \left(2 \cdot (me_1 + me_2 + 0.5) + 0.5\right) \cdot 2^{cb_3},$$

then

Table 3-74. GFLIB_Lut2D_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltIn1,\ fltIn2,\ u32ShamOffset1, \\ u32ShamOffset2,\ \boldsymbol{fltTable} \end{pmatrix} \in X \;\middle\|\; X \cap D \neq \varnothing \wedge X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, or an input is beyond allowed range, i.e. $\left\{ \begin{pmatrix} fltIn1,\ fltIn2,\ u32ShamOffset1, \\ u32ShamOffset2,\ \boldsymbol{fltTable} \end{pmatrix} \in X \;\middle\|\; X \cap M \neq \varnothing \vee \lvert fltIn \rvert \geq 1 \vee \lvert fltIn \rvert \geq 1 \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, inputs within allowed range, i.e. $\left\{ \begin{pmatrix} fltIn1,\ fltIn2,\ u32ShamOffset1, \\ u32ShamOffset2,\ \boldsymbol{fltTable} \end{pmatrix} \in X \;\middle\|\; \begin{matrix} X \subset N \cup \{-0,\ +0\}, \\ \boldsymbol{fltTable} \subset N \cup \{-0,\ +0\} \\ \lvert fltIn1 \rvert < 1, \\ \lvert fltIn2 \rvert < 1, \\ \max_n(\lvert mra_n \rvert) + 2.5 \cdot \mathrm{ULP}(\max_n(mra_n)) \neq Inf, \\ \bigvee mra_n = 0 \vee \lvert mra_n \rvert - 2.5 \cdot \mathrm{ULP}(mra_n) \notin D, \\ \max_n(\lvert sra_n \rvert) + 2.5 \cdot \mathrm{ULP}(\max_n(sra_n)) \neq Inf, \\ \bigvee sra_n = 0 \vee \lvert sra_n \rvert - 2.5 \cdot \mathrm{ULP}(sra_n) \notin D, \\ \max_n(\lvert mrb_n \rvert) + 2.5 \cdot \mathrm{ULP}(\max_n(mrb_n)) \neq Inf, \\ \bigvee mrb_n = 0 \vee \lvert mrb_n \rvert - 2.5 \cdot \mathrm{ULP}(mrb_n) \notin D, \\ \max_n(\lvert srb_n \rvert) + 2.5 \cdot \mathrm{ULP}(\max_n(srb_n)) \neq Inf, \\ \bigvee srb_n = 0 \vee \lvert srb_n \rvert - 2.5 \cdot \mathrm{ULP}(srb_n) \notin D, \\ \max_n(\lvert mrc_n \rvert) + 2.5 \cdot \mathrm{ULP}(\max_n(mrc_n)) \neq Inf, \\ \bigvee mrc_n = 0 \vee \lvert mrc_n \rvert - 2.5 \cdot \mathrm{ULP}(mrc_n) \notin D, \\ refResult = 0 \vee \\ \begin{pmatrix} \lvert refResult \rvert - me_3 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me_3 \cdot \mathrm{ULP}(refResult) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -me_3, me_3 \rangle$	N/A
Normalized or zero input values which cause result underflow, i.e.	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-74. GFLIB_Lut2D_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltIn1, fltIn2, u32ShamOffset1, \\ u32ShamOffset2, \boldsymbol{fltTable} \end{pmatrix} \in X \middle\vert \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert - me_3 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$		
Normalized or zero input values which cause intermediate result underflow, i.e. $\left\{ \begin{pmatrix} fltIn1, fltIn2, u32ShamOffset1, \\ u32ShamOffset2, \boldsymbol{fltTable} \end{pmatrix} \in X \middle\vert \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \exists \lvert mra_n \rvert - 2.5 \cdot \mathrm{ULP}(mra_n) \in D \wedge mra_n \neq 0 \vee \\ \exists \lvert sra_n \rvert - 2.5 \cdot \mathrm{ULP}(sra_n) \in D \wedge sra_n \neq 0 \vee \\ \exists \lvert mrb_n \rvert - 2.5 \cdot \mathrm{ULP}(mrb_n) \in D \wedge mrb_n \neq 0 \vee \\ \exists \lvert srb_n \rvert - 2.5 \cdot \mathrm{ULP}(srb_n) \in D \wedge srb_n \neq 0 \vee \\ \exists \lvert mrc_n \rvert - 2.5 \cdot \mathrm{ULP}(mrc_n) \in D \wedge mrc_n \neq 0 \vee \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}

3.23 Function GFLIB_Ramp_FLT

3.23.1 Declaration

```
tFloat GFLIB_Ramp_FLT(tFloat fltIn, GFLIB_RAMP_T_FLT *const pParam);
```

3.23.2 Arguments

Table 3-75. GFLIB_Ramp_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument representing the desired output value. Input value is in single precision floating data format.
GFLIB_RAMP_T_FLT *const	pParam	**input, output**	Pointer to the ramp parameters structure. Arguments of the structure contain single precision floating point values.

3.23.3 Worst-Case Error Bounds

Let $(fltIn, fltState, fltRampUp, fltRampDown) \in X$ be a vector of inputs to GFLIB_Ramp_FLT,

$fltState$ input represents a real number with an error of max. ±0.5 ulp,

$refResult$ be the theoretical exact result,

$ie_1 = \mathrm{fe}(\,|fltState| + 0.5 \cdot \mathrm{ULP}(fltState)\,)$,

$ie_2 = \mathrm{fe}(refResult)$,

$$cb = \begin{cases} 0, & ie_1 - ie_2 \le 0 \\ ie_1 - ie_2, & ie_1 - ie_2 > 0 \end{cases},$$

$me = 0.5 + 0.5 \cdot 2^{cb}$,

then

Table 3-76. GFLIB_Ramp_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltIn, fltState, \\ fltRampUp, fltRampDown \end{pmatrix} \in X \;\middle|\; X \cap D \neq \varnothing \wedge X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltIn, fltState, \\ fltRampUp, fltRampDown \end{pmatrix} \in X \;\middle|\; X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltIn, fltState, \\ fltRampUp, fltRampDown \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \left(\begin{matrix} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{matrix} \right) \end{matrix} \right\}$	$\langle -me, me \rangle$	N/A
Normalized or zero input values which cause result underflow, i.e.	$(-Inf, Inf)$	N/A

Table 3-76. GFLIB_Ramp_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{pmatrix} fltIn,\ fltState, \\ fltRampUp,\ fltRampDown \end{pmatrix} \in X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$		

Table 3-77. GFLIB_Ramp_FLT Worst-Case Error Bounds - fltState Output

Subset of input domain	Worst-case error bounds for fltState output [ulp]	Allowed specific values (regardless the error bounds) for fltState output
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{pmatrix} fltIn,\ fltState, \\ fltRampUp,\ fltRampDown \end{pmatrix} \in X \;\middle|\; X \cap D \neq \emptyset \wedge X \cap M = \emptyset \right\}$	$(-Inf,\ Inf)$	N/A
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} fltIn,\ fltState, \\ fltRampUp,\ fltRampDown \end{pmatrix} \in X \;\middle|\; X \cap M \neq \emptyset \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltIn,\ fltState, \\ fltRampUp,\ fltRampDown \end{pmatrix} \in X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + me \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -me,\ me \rangle$	N/A
Normalized or zero input values which cause result underflow, i.e. $\left\{ \begin{pmatrix} fltIn,\ fltState, \\ fltRampUp,\ fltRampDown \end{pmatrix} \in X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0,\ +0\}, \\ \lvert refResult \rvert - me \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$(-Inf,\ Inf)$	N/A

Table 3-78. GFLIB_Ramp_FLT Worst-Case Error Bounds - fltInK1 Output

Subset of input domain	Worst-case error bounds for fltInK1 output [ulp]	Allowed specific values (regardless the error bounds) for fltInK1 output
Entire input domain	0	*fltIn*

3.24 Function GFLIB_Sign_FLT

3.24.1 Declaration

```
tFloat GFLIB_Sign_FLT(tFloat fltIn);
```

3.24.2 Arguments

Table 3-79. GFLIB_Sign_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument.

3.24.3 Worst-Case Error Bounds

Let $fltIn \in X$ be an input to GFLIB_Sign_FLT,

then

Table 3-80. GFLIB_Sign_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input NaN, i.e. $\{fltIn \in X \mid fltIn = NaN\}$	N/A	{NaN, *pmax*, *nmax*, 0, -0}
Normalized, denormalized, infinite, or zero input, i.e. $\{fltIn \in X \mid fltIn \neq NaN\}$	0	N/A

3.25 Function GFLIB_Sin_FLT

3.25.1 Declaration

```
tFloat GFLIB_Sin_FLT(tFloat fltIn, const GFLIB_SIN_T_FLT *const pParam);
```

3.25.2 Arguments

Table 3-81. GFLIB_Sin_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument is a single precision floating point number that contains an angle in radians between [- π, π].
const GFLIB_SIN_T_FLT *const	pParam	**input**	Pointer to an array of approximation coefficients.

3.25.3 Worst-Case Error Bounds

Only the default approximation coefficients are considered.

Let $(fltIn) \in X$ be an input to GFLIB_Sin_FLT,

π_{SP} be the single-precision representation of π,

then

Table 3-82. GFLIB_Sin_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input infinite, i.e. $\{fltIn \in X \mid fltIn \in \{Inf, -Inf\}\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Input NaN, i.e. $\{fltIn \in X \mid fltIn = NaN\}$	$(-Inf, Inf)$	{NaN, pmax}
Normalized, denormalized, or zero input within guaranteed range, i.e. $\{fltIn \in X \mid fltIn \notin M, \lvert fltIn \rvert \le \pi_{SP}\}$	$\langle -334, 334 \rangle$	N/A

Table continues on the next page...

Table 3-82. GFLIB_Sin_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Normalized input beyond the guaranteed range, i.e. $\{fltIn \in X \mid fltIn \notin M, \lvert fltIn \rvert > \pi_{SP}\}$	$(-Inf, Inf)$	N/A

3.26 Function GFLIB_SinCos_FLT

3.26.1 Declaration

```
void GFLIB_SinCos_FLT(tFloat fltIn, SWLIBS_2Syst_FLT *pOut, const GFLIB_SINCOS_T_FLT *const
pParam);
```

3.26.2 Arguments

Table 3-83. GFLIB_SinCos_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument is a single precision floating point number that contains an angle in radians between [- π, π].
SWLIBS_2Syst_FLT *	pOut	**output**	Pointer to the structure where the values of the sine and cosine of the input angle are stored. The function returns the sine and cosine of the input argument as a single precision floating point number. The *sine* of input angle is returned in first item of the structure and the *cosine* of input angle is returned in second item of the structure.
const GFLIB_SINCOS_T_FLT *const	pParam	**input**	Pointer to an array of approximation coefficients.

3.26.3 Worst-Case Error Bounds

Only the default approximation coefficients are considered.

Let $(fltIn) \in X$ be an input to GFLIB_SinCos_FLT,

π_{SP} be the single-precision representation of π,

then

Table 3-84. GFLIB_SinCos_FLT Worst-Case Error Bounds - fltArg1 Output

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
Input infinite, i.e. $\{ fltIn \in X \mid fltIn \in \{ Inf, -Inf \} \}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Input NaN, i.e. $\{ fltIn \in X \mid fltIn = NaN \}$	$(-Inf, Inf)$	{NaN, pmax}
Normalized, denormalized, or zero input within guaranteed range, i.e. $\{ fltIn \in X \mid fltIn \notin M, \lvert fltIn \rvert \le \pi_{SP} \}$	$\langle -334, 334 \rangle$	N/A
Normalized input beyond the guaranteed range, i.e. $\{ fltIn \in X \mid fltIn \notin M, \lvert fltIn \rvert > \pi_{SP} \}$	$(-Inf, Inf)$	N/A

Table 3-85. GFLIB_SinCos_FLT Worst-Case Error Bounds - fltArg2 Output

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
Input NaN or infinity, i.e. $\{ fltIn \in X \mid fltIn \in M \}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Input denormalized, normalized or zero, and within the allowed input interval, i.e. $\left\{ \begin{matrix} fltIn \in X \mid X \cap M = \varnothing, \\ -\pi \le fltIn \le \pi \end{matrix} \right\}$	$\langle -334, 334 \rangle$	N/A
Input normalized and beyond the allowed range, i.e. $\left\{ \begin{matrix} fltIn \in X \mid X \cap M = \varnothing, \\ fltIn < -\pi \vee fltIn > \pi \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, -Inf, Inf}

3.27 Function GFLIB_Sqrt_FLT

3.27.1 Declaration

```
tFloat GFLIB_Sqrt_FLT(tFloat fltIn);
```

3.27.2 Arguments

Table 3-86. GFLIB_Sqrt_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	The input value.

3.27.3 Worst-Case Error Bounds

Let $fltIn \in X$ be an input to GFLIB_Sqrt_FLT,

then

Table 3-87. GFLIB_Sqrt_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input positive infinity, i.e. $\{ fltIn \in X \mid fltIn = Inf \}$	$(-Inf, Inf)$	{NaN, Inf, -Inf, *pmax*, 0}
Input NaN, i.e. $\{ fltIn \in X \mid fltIn = NaN \}$	$(-Inf, Inf)$	{0, -0, NaN}
Negative input, i.e. $\{ fltIn \in X \mid fltIn < 0 \}$	N/A	{-0, NaN}
Positive denormalized input, i.e. $\{ fltIn \in X \mid fltIn \in D \wedge fltIn > 0 \}$	$(-Inf, Inf)$	N/A
Positive normalized input, i.e. $\{ fltIn \in X \mid fltIn \in N \wedge fltIn > 0 \}$	$\langle -0.5, 0.5 \rangle$	N/A
Positive zero input, i.e. $\{ fltIn \in X \mid fltIn = 0 \}$	N/A	0
Negative zero input, i.e. $\{ fltIn \in X \mid fltIn = -0 \}$	N/A	-0

3.28 Function GFLIB_Tan_FLT

3.28.1 Declaration

```
tFloat GFLIB_Tan_FLT(tFloat fltIn, const GFLIB_TAN_T_FLT *const pParam);
```

3.28.2 Arguments

Table 3-88. GFLIB_Tan_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input argument is a single precision floating point number that contains an angle in radians between (- π, π).
const GFLIB_TAN_T_FLT *const	pParam	**input**	Pointer to an array of approximation coefficients.

3.28.3 Worst-Case Error Bounds

Only the default approximation coefficients are considered.

Let $fltIn \in X$ be an input to GFLIB_Tan_FLT,

π_{SP} be the single-precision representation of π,

then

Table 3-89. GFLIB_Tan_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input infinite, i.e. $\{fltIn \in X \mid fltIn \in \{Inf, -Inf\}\}$	$(-Inf, Inf)$	{Inf, -Inf, NaN}
Input NaN, i.e. $\{fltIn \in X \mid fltIn = NaN\}$	$(-Inf, Inf)$	NaN
Normalized, denormalized, or zero input within guaranteed range, i.e. $\{fltIn \in X \mid fltIn \notin M, \lvert fltIn \rvert \leq \pi_{SP}\}$	$\langle -13, 13 \rangle$	N/A
Normalized input beyond the guaranteed range, i.e. $\{fltIn \in X \mid fltIn \notin M, \lvert fltIn \rvert > \pi_{SP}\}$	$(-Inf, Inf)$	N/A

3.29 Function GFLIB_UpperLimit_FLT

3.29.1 Declaration

```
tFloat GFLIB_UpperLimit_FLT(tFloat fltIn, const GFLIB_UPPERLIMIT_T_FLT *const pParam);
```

3.29.2 Arguments

Table 3-90. GFLIB_UpperLimit_FLT arguments

Type	Name	Direction	Description
tFloat	fltIn	**input**	Input value.
const GFLIB_UPPERLIMIT_T_FLT *const	pParam	**input**	Pointer to the limits structure.

3.29.3 Worst-Case Error Bounds

Let $(fltIn, fltUpperLimit) \in X$ be a vector of inputs to GFLIB_UpperLimit_FLT,

then

Table 3-91. GFLIB_UpperLimit_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs are normalized, denormalized, i.e. $\left\{ (fltIn, fltUpperLimit) \in X \;\middle|\; X \cap M = \varnothing \right\}$	0	N/A
At least one of the inputs is NaN or infinite, i.e. $\left\{ (fltIn, fltUpperLimit) \in X \;\middle|\; X \cap M \neq \varnothing \vee \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

3.30 Function GFLIB_VectorLimit_FLT

3.30.1 Declaration

```
tBool GFLIB_VectorLimit_FLT(SWLIBS_2Syst_FLT *const pOut, const SWLIBS_2Syst_FLT *const pIn,
const GFLIB_VECTORLIMIT_T_FLT *const pParam);
```

3.30.2 Arguments

Table 3-92. GFLIB_VectorLimit_FLT arguments

Type	Name	Direction	Description
const SWLIBS_2Syst_FLT *const	pIn	**input**	Pointer to the structure of the input vector.
SWLIBS_2Syst_FLT *const	pOut	**output**	Pointer to the structure of the limited output vector.
const GFLIB_VECTORLIMIT_T_FLT *const	pParam	**input**	Pointer to the parameters structure.

3.30.3 Worst-Case Error Bounds

Let $(fltArg1, fltArg2, fltLimit) \in X$ be a vector of inputs to GFLIB_VectorLimit_FLT, $reffltArg1$, $reffltArg2$ be the theoretical exact results for outputs $fltArg1$ and $fltArg2$, respectively,

$ir_1 = fltArg1 \cdot fltArg1$,

$ir_2 = fltArg2 \cdot fltArg2$,

$ir_3 = ir_1 + ir_2$,

$ir_4 = \sqrt{ir_3}$,

$ir_5 = \frac{fltLimit}{ir_4}$,

$ir_6 = fltArg1 \cdot ir_5$,

$ir_7 = fltArg2 \cdot ir_5$,

then

Table 3-93. GFLIB_VectorLimit_FLT Return Value

Subset of input domain	Meaningful Return Value
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2,\ fltLimit) \in X \ \middle| \\ X \cap D \neq \emptyset \wedge X \cap M = \emptyset \end{array} \right\}$	No
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2,\ fltLimit) \in X \ \middle| \\ X \cap M \neq \emptyset \end{array} \right\}$	No
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2,\ fltLimit) \in X \ \middle| \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{array}{c} reffltArg1,\ reffltArg2, \\ ir_1,\ ir_2,\ ir_3,\ ir_4,\ ir_5,\ ir_6,\ ir_7 \end{array} \right\} \subset N \cup \{-0,\ +0\} \end{array} \right\}$	Yes
Normalized or zero input values which cause an output underflow, i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2,\ fltLimit) \in X \ \middle| \\ X \subset N \cup \{-0,\ +0\}, \\ \{reffltArg1,\ reffltArg2\} \cap M \neq \emptyset \\ \left(\begin{array}{c} \lvert reffltArg1 \rvert - 276 \cdot \mathrm{ULP}(reffltArg1) \in D \wedge \\ reffltArg1 \neq 0 \end{array} \right) \vee \\ \left(\begin{array}{c} \lvert reffltArg2 \rvert - 276 \cdot \mathrm{ULP}(reffltArg2) \in D \wedge \\ reffltArg2 \neq 0 \end{array} \right) \end{array} \right\}$	No
Normalized or zero input values which cause an output overflow, i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2,\ fltLimit) \in X \ \middle| \\ X \subset N \cup \{-0,\ +0\}, \\ \left(\lvert reffltArg1 \rvert + 276 \cdot \mathrm{ULP}(reffltArg1) \in M \wedge \right) \vee \\ \left(\lvert reffltArg2 \rvert + 276 \cdot \mathrm{ULP}(reffltArg2) \in M \wedge \right) \end{array} \right\}$	No
Normalized or zero input values which cause intermediate results underflow, i.e.	No

Table continues on the next page...

Table 3-93. GFLIB_VectorLimit_FLT Return Value (continued)

Subset of input domain	Meaningful Return Value
$$\left\{\begin{matrix} \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \\ X \subset N \cup \{-0, +0\}, \\ (\lvert ir_1\rvert - 0.5 \cdot \mathrm{ULP}(ir_1) \in D \wedge ir_1 \neq 0) \vee \\ (\lvert ir_2\rvert - 0.5 \cdot \mathrm{ULP}(ir_2) \in D \wedge ir_2 \neq 0) \vee \\ (\lvert ir_3\rvert - \mathrm{ULP}(ir_3) \in D \wedge ir_3 \neq 0) \vee \\ (\lvert ir_4\rvert - 276 \cdot \mathrm{ULP}(ir_4) \in D \wedge ir_4 \neq 0) \vee \\ (\lvert ir_5\rvert - 276 \cdot \mathrm{ULP}(ir_5) \in D \wedge ir_5 \neq 0) \vee \\ (\lvert ir_6\rvert - 276 \cdot \mathrm{ULP}(ir_6) \in D \wedge ir_6 \neq 0) \vee \\ (\lvert ir_7\rvert - 276 \cdot \mathrm{ULP}(ir_7) \in D \wedge ir_7 \neq 0), \\ \{ir_1, ir_2, ir_3, ir_4, ir_5, ir_6, ir_7\} \cap M \neq \varnothing \end{matrix}\right\}$$	
Normalized or zero input values which cause intermediate results overflow, i.e. $$\left\{\begin{matrix} \begin{pmatrix} fltInErr, fltPropGain, fltIntegGain, \\ fltIntegPartK\_1, fltInK\_1 \end{pmatrix} \in X \\ X \subset N \cup \{-0, +0\}, \\ (\lvert ir_1\rvert + 0.5 \cdot \mathrm{ULP}(ir_1) \in M) \vee \\ (\lvert ir_2\rvert + 0.5 \cdot \mathrm{ULP}(ir_2) \in M) \vee \\ (\lvert ir_3\rvert + \mathrm{ULP}(ir_3) \in M) \vee \\ (\lvert ir_4\rvert + 276 \cdot \mathrm{ULP}(ir_4) \in M) \vee \\ (\lvert ir_5\rvert + 276 \cdot \mathrm{ULP}(ir_5) \in M) \vee \\ (\lvert ir_6\rvert + 276 \cdot \mathrm{ULP}(ir_6) \in M) \vee \\ (\lvert ir_7\rvert + 276 \cdot \mathrm{ULP}(ir_7) \in M) \end{matrix}\right\}$$	No

Table 3-94. GFLIB_VectorLimit_FLT Worst-Case Error Bounds - fltArg1 Output

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
Any of the inputs is denormalized, others are normalized or zero, i.e. $$\left\{\begin{matrix} (fltArg1, fltArg2, fltLimit) \in X \\ X \cap D \neq \varnothing \wedge X \cap M = \varnothing \end{matrix}\right\}$$	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-94. GFLIB_VectorLimit_FLT Worst-Case Error Bounds - fltArg1 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
Any of the inputs is NaN or infinity, i.e. $\left\{ (fltArg1,\ fltArg2,\ fltLimit) \in X \;\middle|\; X \cap M \neq \varnothing \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{matrix} (fltArg1,\ fltArg2,\ fltLimit) \in X \;\vert \\ X \subset N \cup \{-0,\ +0\}, \\ \left\{ \begin{matrix} reffltArg1, reffltArg2, \\ ir_1, ir_2, ir_3, ir_4, ir_5, ir_6, ir_7 \end{matrix} \right\} \subset N \cup \{-0,\ +0\} \end{matrix} \right\}$	$\langle -276, 276 \rangle$	N/A
Normalized or zero input values which cause an output underflow, i.e. $\left\{ \begin{matrix} (fltArg1,\ fltArg2,\ fltLimit) \in X \;\vert \\ X \subset N \cup \{-0,\ +0\}, \\ \{reffltArg1, reffltArg2\} \cap M \neq \varnothing \\ \left(\begin{matrix} \vert reffltArg1\vert - 276 \cdot \mathrm{ULP}(reffltArg1) \in D \wedge \\ reffltArg1 \neq 0 \end{matrix} \right) \vee \\ \left(\begin{matrix} \vert reffltArg2\vert - 276 \cdot \mathrm{ULP}(reffltArg2) \in D \wedge \\ reffltArg2 \neq 0 \end{matrix} \right) \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause an output overflow, i.e. $\left\{ \begin{matrix} (fltArg1,\ fltArg2,\ fltLimit) \in X \;\vert \\ X \subset N \cup \{-0,\ +0\}, \\ (\vert reffltArg1\vert + 276 \cdot \mathrm{ULP}(reffltArg1) \in M \wedge\ ) \vee \\ (\vert reffltArg2\vert + 276 \cdot \mathrm{ULP}(reffltArg2) \in M \wedge\ ) \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate results underflow, i.e.	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table continues on the next page...

Table 3-94. GFLIB_VectorLimit_FLT Worst-Case Error Bounds - fltArg1 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
$\left\{ \begin{array}{c} \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \ \middle\vert \\ X \subset N \cup \{-0,\ +0\}, \\ (\lvert ir_1 \rvert - 0.5 \cdot \mathrm{ULP}(ir_1) \in D \wedge ir_1 \neq 0) \vee \\ (\lvert ir_2 \rvert - 0.5 \cdot \mathrm{ULP}(ir_2) \in D \wedge ir_2 \neq 0) \vee \\ (\lvert ir_3 \rvert - \mathrm{ULP}(ir_3) \in D \wedge ir_3 \neq 0) \vee \\ (\lvert ir_4 \rvert - 276 \cdot \mathrm{ULP}(ir_4) \in D \wedge ir_4 \neq 0) \vee \\ (\lvert ir_5 \rvert - 276 \cdot \mathrm{ULP}(ir_5) \in D \wedge ir_5 \neq 0) \vee \\ (\lvert ir_6 \rvert - 276 \cdot \mathrm{ULP}(ir_6) \in D \wedge ir_6 \neq 0) \vee \\ (\lvert ir_7 \rvert - 276 \cdot \mathrm{ULP}(ir_7) \in D \wedge ir_7 \neq 0), \\ \{ir_1,\ ir_2,\ ir_3,\ ir_4,\ ir_5,\ ir_6,\ ir_7\} \cap M \neq \emptyset \end{array} \right\}$		
Normalized or zero input values which cause intermediate results overflow, i.e. $\left\{ \begin{array}{c} \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \ \middle\vert \\ X \subset N \cup \{-0,\ +0\}, \\ (\lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1) \in M) \vee \\ (\lvert ir_2 \rvert + 0.5 \cdot \mathrm{ULP}(ir_2) \in M) \vee \\ (\lvert ir_3 \rvert + \mathrm{ULP}(ir_3) \in M) \vee \\ (\lvert ir_4 \rvert + 276 \cdot \mathrm{ULP}(ir_4) \in M) \vee \\ (\lvert ir_5 \rvert + 276 \cdot \mathrm{ULP}(ir_5) \in M) \vee \\ (\lvert ir_6 \rvert + 276 \cdot \mathrm{ULP}(ir_6) \in M) \vee \\ (\lvert ir_7 \rvert + 276 \cdot \mathrm{ULP}(ir_7) \in M) \end{array} \right\}$	$(-Inf,\ Inf)$	{Inf, –Inf, NaN}

Table 3-95. GFLIB_VectorLimit_FLT Worst-Case Error Bounds - fltArg2 Output

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
Any of the inputs is denormalized, others are normalized or zero, i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2,\ fltLimit) \in X \ \middle\vert \\ X \cap D \neq \emptyset \wedge X \cap M = \emptyset \end{array} \right\}$	$(-Inf,\ Inf)$	N/A

Table continues on the next page...

Table 3-95. GFLIB_VectorLimit_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltLimit) \in X \mid \\ X \cap M \neq \varnothing \end{array} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
Normalized or zero input values which lead to normalized or zero results, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltLimit) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \left\{ \begin{array}{c} reffltArg1, reffltArg2, \\ ir_1, ir_2, ir_3, ir_4, ir_5, ir_6, ir_7 \end{array} \right\} \subset N \cup \{-0, +0\} \end{array} \right\}$	$\langle -276, 276 \rangle$	N/A
Normalized or zero input values which cause an output underflow, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltLimit) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \{reffltArg1, reffltArg2\} \cap M \neq \varnothing \\ \left(\begin{array}{c} \lvert reffltArg1 \rvert - 276 \cdot \mathrm{ULP}(reffltArg1) \in D \wedge \\ reffltArg1 \neq 0 \end{array} \right) \vee \\ \left(\begin{array}{c} \lvert reffltArg2 \rvert - 276 \cdot \mathrm{ULP}(reffltArg2) \in D \wedge \\ reffltArg2 \neq 0 \end{array} \right) \end{array} \right\}$	$(-Inf, Inf)$	N/A
Normalized or zero input values which cause an output overflow, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltLimit) \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ (\lvert reffltArg1 \rvert + 276 \cdot \mathrm{ULP}(reffltArg1) \in M \wedge) \vee \\ (\lvert reffltArg2 \rvert + 276 \cdot \mathrm{ULP}(reffltArg2) \in M \wedge) \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}
Normalized or zero input values which cause intermediate results underflow, i.e.	$(-Inf, Inf)$	{Inf, –Inf, NaN}

Table continues on the next page...

Table 3-95. GFLIB_VectorLimit_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{ \begin{array}{c} \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \\ X \subset N \cup \{-0,\ +0\}, \\ (\lvert ir_1 \rvert - 0.5 \cdot \mathrm{ULP}(ir_1) \in D \wedge ir_1 \neq 0) \vee \\ (\lvert ir_2 \rvert - 0.5 \cdot \mathrm{ULP}(ir_2) \in D \wedge ir_2 \neq 0) \vee \\ (\lvert ir_3 \rvert - \mathrm{ULP}(ir_3) \in D \wedge ir_3 \neq 0) \vee \\ (\lvert ir_4 \rvert - 276 \cdot \mathrm{ULP}(ir_4) \in D \wedge ir_4 \neq 0) \vee \\ (\lvert ir_5 \rvert - 276 \cdot \mathrm{ULP}(ir_5) \in D \wedge ir_5 \neq 0) \vee \\ (\lvert ir_6 \rvert - 276 \cdot \mathrm{ULP}(ir_6) \in D \wedge ir_6 \neq 0) \vee \\ (\lvert ir_7 \rvert - 276 \cdot \mathrm{ULP}(ir_7) \in D \wedge ir_7 \neq 0), \\ \{ir_1, ir_2, ir_3, ir_4, ir_5, ir_6, ir_7\} \cap M \neq \varnothing \end{array} \right\}$		
Normalized or zero input values which cause intermediate results overflow, i.e. $\left\{ \begin{array}{c} \left(\begin{array}{c} fltInErr,\ fltPropGain,\ fltIntegGain, \\ fltIntegPartK\_1,\ fltInK\_1 \end{array} \right) \in X \\ X \subset N \cup \{-0,\ +0\}, \\ (\lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1) \in M) \vee \\ (\lvert ir_2 \rvert + 0.5 \cdot \mathrm{ULP}(ir_2) \in M) \vee \\ (\lvert ir_3 \rvert + \mathrm{ULP}(ir_3) \in M) \vee \\ (\lvert ir_4 \rvert + 276 \cdot \mathrm{ULP}(ir_4) \in M) \vee \\ (\lvert ir_5 \rvert + 276 \cdot \mathrm{ULP}(ir_5) \in M) \vee \\ (\lvert ir_6 \rvert + 276 \cdot \mathrm{ULP}(ir_6) \in M) \vee \\ (\lvert ir_7 \rvert + 276 \cdot \mathrm{ULP}(ir_7) \in M) \end{array} \right\}$	$(-Inf, Inf)$	{Inf, –Inf, NaN}

3.31 Function GMCLIB_Clark_FLT

3.31.1 Declaration

```
void GMCLIB_Clark_FLT(SWLIBS_2Syst_FLT *const pOut, const SWLIBS_3Syst_FLT *const pIn);
```

3.31.2 Arguments

Table 3-96. GMCLIB_Clark_FLT arguments

Type	Name	Direction	Description
const SWLIBS_3Syst_FLT *const	pIn	**input**	Pointer to the structure containing data of the three-phase stationary system (fltA-fltB-fltC). Arguments of the structure contain single precision floating point values.
SWLIBS_2Syst_FLT *const	pOut	**output**	Pointer to the structure containing data of the two-phase stationary orthogonal system (α- β). Arguments of the structure contain single precision floating point values.

3.31.3 Worst-Case Error Bounds

Let $(fltArg1, fltArg2, fltArg3) \in X$ be a vector of inputs to GMCLIB_Clark_FLT,

$reffltArg1$, $reffltArg2$ be the theoretical exact results for outputs $fltArg1$ and $fltArg2$, respectively,

$ir = fltArg2 - fltArg3$,

then

Table 3-97. GMCLIB_Clark_FLT Worst-Case Error Bounds - fltArg1 Output

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Entire input domain	0	*fltArg1*

Table 3-98. GMCLIB_Clark_FLT Worst-Case Error Bounds - fltArg2 Output

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
Both inputs infinity, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 \in \{Inf, -Inf\}, \\ fltArg3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*}
fltArg2 arbitrary value, *fltArg3* NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg3 = NaN \end{array} \right\}$	N/A	{NaN, *pmax*, *nmax*}

Table continues on the next page...

Table 3-98. GMCLIB_Clark_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
fltArg2 infinity, *fltArg3* normalized, denormalized, or zero, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 \in \{Inf, -Inf\}, \\ fltArg3 \in N \cup D \cup \{-0, +0\}\end{matrix}\right\}$	N/A	{sign(*fltArg2*)•Inf, sign(*fltArg2*)•pmax}
fltArg2 NaN, *fltArg3* arbitrary value, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 = NaN\end{matrix}\right\}$	N/A	{NaN, *pmax*, *nmax*}
fltArg2 normalized, denormalized, or zero, *fltArg3* infinity, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 \in N \cup D \cup \{-0, +0\}, \\ fltArg3 \in \{Inf, -Inf\}\end{matrix}\right\}$	N/A	{-sign(*fltArg3*)•Inf, -sign(*fltArg3*)•pmax}
Both inputs denormalized, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ X \subset D\end{matrix}\right\}$	$(-Inf, Inf)$	{-0, +0}
fltArg2 denormalized, *fltArg3* zero, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 \in D, fltIn_2 = 0\end{matrix}\right\}$	$(-Inf, Inf)$	{-0, +0}
fltArg2 denormalized, *fltArg3* normalized, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 \in D, \\ fltArg3 \in N\end{matrix}\right\}$	$(-Inf, Inf)$	N/A
fltArg2 zero, *fltArg3* denormalized, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 = 0, fltArg3 \in D\end{matrix}\right\}$	$(-Inf, Inf)$	{-0, +0}
Both inputs zero, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 = 0, fltArg3 = 0\end{matrix}\right\}$	0	N/A
fltArg2 zero, *fltArg3* normalized, i.e. $\left\{\begin{matrix}(fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 = 0, \\ fltArg3 \in N\end{matrix}\right\}$	$\langle -2.5, 2.5 \rangle$	N/A
fltArg2 normalized, *fltArg3* denormalized, i.e.	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-98. GMCLIB_Clark_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 \in N, fltArg3 \in D \end{array} \right\}$		
fltArg2 normalized, *fltArg3* zero, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ fltArg2 \in N, fltArg3 = 0 \end{array} \right\}$	$\langle -2.5, 2.5 \rangle$	N/A
Both inputs normalized, results normalized or zero, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ X \subset N, \\ ref fltArg2 = 0 \vee \\ \left(\begin{array}{c} \lvert ref fltArg2 \rvert - 0.5 \cdot \mathrm{ULP}(ref fltArg2) \notin D \wedge \\ \lvert ref fltArg2 \rvert + 0.5 \cdot \mathrm{ULP}(ref fltArg2) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -2.5, 2.5 \rangle$	N/A
Both inputs normalized, result overflow, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ X \subset N, \\ \lvert ref fltArg2 \rvert + 0.5 \cdot \mathrm{ULP}(ref fltArg2) = Inf \end{array} \right\}$	$\langle -2.5, 2.5 \rangle$	{sign(reffltArg2)•Inf, sign(reffltArg2)•pmax}
Both inputs normalized, intermediate result overflow, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ X \subset N, \\ \lvert ir \rvert + 0.5 \cdot \mathrm{ULP}(ir) = Inf \end{array} \right\}$	$\langle -2.5, 2.5 \rangle$	{sign(reffltArg2)•Inf, sign(reffltArg2)•pmax}
Both inputs normalized, intermediate result underflow, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2, fltArg3) \in X \mid \\ X \subset N, \\ \lvert ir \rvert - 0.5 \cdot \mathrm{ULP}(ir) \in D \wedge \\ ir \neq 0 \end{array} \right\}$	$\langle -2.5, 2.5 \rangle$	{-0, +0}

3.32 Function GMCLIB_ClarkInv_FLT

3.32.1 Declaration

```
void GMCLIB_ClarkInv_FLT(SWLIBS_3Syst_FLT *const pOut, const SWLIBS_2Syst_FLT *const pIn);
```

3.32.2 Arguments

Table 3-99. GMCLIB_ClarkInv_FLT arguments

Type	Name	Direction	Description
const SWLIBS_2Syst_FLT *const	pIn	**input**	Pointer to the structure containing data of the two-phase stationary orthogonal system (α- β). Arguments of the structure contain single precision floating point values.
SWLIBS_3Syst_FLT *const	pOut	**output**	Pointer to the structure containing data of the three-phase stationary system (fltA-fltB-fltC). Arguments of the structure contain single precision floating point values.

3.32.3 Worst-Case Error Bounds

Let $(fltArg1, fltArg2) \in X$ be a vector of inputs to GMCLIB_ClarkInv_FLT,

$reffltArg2$, $reffltArg3$ be the theoretical exact results for outputs $fltArg2$ and $fltArg3$, respectively,

$$ir_1 = 0.5 \cdot fltArg1,$$

$$ir_2 = \frac{\sqrt{3}}{2} \cdot fltArg2,$$

$$ie_1 = \max(\mathrm{fe}(|ir_1|), \mathrm{fe}(|ir_2| + 1.5 \cdot \mathrm{ULP}(ir_2))),$$

$$refInteg = \ fltIntegPartK\_1 + ir_2,$$

$$ix = \mathrm{fe}(reffltArg2),$$

$$iy = \mathrm{fe}(reffltArg3),$$

$$cb_1 = \begin{cases} 0, & ie_1 - ix \le 0 \\ ie_1 - ix, & ie_1 - ix > 0 \end{cases},$$

$$cb_2 = \begin{cases} 0, & ie_1 - iy \le 0 \\ ie_1 - iy, & ie_1 - iy > 0 \end{cases},$$

$$me_1 = 0.5 + 1.5 \cdot 2^{cb_1},$$

$$me_2 = 0.5 + 1.5 \cdot 2^{cb_2},$$

then

Table 3-100. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg1*

Subset of input domain	Worst-case error bounds for *fltArg1* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg1* output
Entire input domain	0	N/A

Table 3-101. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg2*

Subset of input domain	Worst-case error bounds for *fltArg2* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg2* output
Both inputs infinity, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ X \subset \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0} if sign(*fltArg1*)=sign(*fltArg2*), {sign(*fltArg2*)•Inf, sign(*fltArg2*)•pmax} otherwise
fltArg1 infinity, *fltArg2* NaN, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = NaN \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0, *pmax*, *nmax*}
fltArg1 infinity, *fltArg2* denormalized, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D \end{matrix} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg1*)•Inf, sign(*fltArg1*)•pmax}
fltArg1 infinity, *fltArg2* zero, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0 \end{matrix} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg1*)•Inf, sign(*fltArg1*)•pmax}
fltArg1 infinity, *fltArg2* normalized, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N \end{matrix} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg1*)•Inf, sign(*fltArg1*)•pmax}
fltArg1 NaN, *fltArg2* infinity, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = NaN, \\ fltIn_2 \in \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
Both inputs NaN, i.e.		{NaN, *pmax*, *nmax*}

Table continues on the next page...

Table 3-101. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg2* (continued)

Subset of input domain	Worst-case error bounds for *fltArg2* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg2* output
$\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 = NaN, fltIn_2 = NaN \right\}$	$(-Inf, Inf)$	
fltArg1 NaN, *fltArg2* denormalized, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 = NaN, fltIn_2 \in D \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
fltArg1 NaN, *fltArg2* zero, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 = NaN, fltIn_2 \in D \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
fltArg1 NaN, *fltArg2* normalized, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 = NaN, fltIn_2 \in N \right\}$	$(-Inf, Inf)$	NaN
fltArg1 denormalized, *fltArg2* infinite, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 \in D, fltIn_2 \in \{Inf, -Inf\} \right\}$	$(-Inf, Inf)$	{sign(*fltArg2*)•Inf, sign(*fltArg2*)•pmax}
fltArg1 denormalized, *fltArg2* NaN, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 \in D, fltIn_2 = NaN \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
Both inputs denormalized, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 \in D, fltIn_2 \in D \right\}$	$\langle -2^{23}, 2^{23} \rangle$	{0, -0}
fltArg1 zero, *fltArg2* infinite, i.e.	$(-Inf, Inf)$	{sign(*fltArg2*)•Inf, sign(*fltArg2*)•pmax}

Table continues on the next page...

Table 3-101. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg2* (continued)

Subset of input domain	Worst-case error bounds for *fltArg2* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg2* output
$\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 = 0, fltIn_2 \in \{Inf, -Inf\} \right\}$		
fltArg1 zero, *fltArg2* NaN, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 = 0, fltIn_2 = NaN \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
fltArg1 denormalized, *fltArg2* zero, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 \in D, fltIn_2 = 0 \right\}$	$\langle -0.5, 0.5 \rangle$	{0, -0}
fltArg1 zero, *fltArg2* denormalized, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 = 0, fltIn_2 \in D \right\}$	$\langle -1.5, 1.5 \rangle$	{0, -0}
fltArg1 normalized, *fltArg2* infinite, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 \in N, fltIn_2 \in \{Inf, -Inf\} \right\}$	$(-Inf, Inf)$	sign(*fltArg2*)•Infsign(*fltArg* 2)•Inf
fltArg1 normalized, *fltArg2* NaN, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 \in N, fltIn_2 = NaN \right\}$	$(-Inf, Inf)$	NaN
fltArg1 denormalized, *fltArg2* normalized, i.e. $\left\{ (fltArg1, fltArg2) \in X \;\middle|\; fltIn_1 \in D, fltIn_2 \in N \right\}$	$\langle -2^{23}, 2^{23} \rangle$	N/A
fltArg1 normalized, *fltArg2* denormalized, i.e.	$\langle -2^{23}, 2^{23} \rangle$	N/A

Table continues on the next page...

Table 3-101. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg2* (continued)

Subset of input domain	Worst-case error bounds for *fltArg2* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg2* output
$\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N \end{matrix} \right\}$		
Both inputs normalized or zero, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ X \in N \cup \{-0, +0\} \end{matrix} \right\}$	$\langle -me_1, me_1 \rangle$	N/A
Both inputs normalized or zero, *fltArg2* output underflow i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ X \in N \cup \{-0, +0\}, \\ \left(\begin{matrix} \lvert ref fltArg2 \rvert - me_1 \cdot \mathrm{ULP}(ref fltArg2) \in D \wedge \\ ref fltArg2 \neq 0, \end{matrix} \right) \end{matrix} \right\}$	$(-Inf, Inf)$	{-0, +0}
Both inputs normalized or zero, intermediate result underflow i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ X \in N \cup \{-0, +0\}, \\ ir_1 \in D \vee \\ \left(\begin{matrix} \lvert ir_2 \rvert - 1.5 \cdot \mathrm{ULP}(ir_2) \in D \wedge \\ ir_2 \neq 0, \end{matrix} \right) \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, -Inf, Inf}
Both inputs normalized or zero, *fltArg2* output overflow i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ X \in N \cup \{-0, +0\}, \\ \lvert ref fltArg2 \rvert + me_1 \cdot \mathrm{ULP}(ref fltArg2) \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, -Inf, Inf}

Table 3-102. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg3*

Subset of input domain	Worst-case error bounds for fltArg3 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg3 output
Both inputs infinity, i.e. $\left\{ \begin{matrix} (fltArg1, fltArg2) \in X \mid \\ X \subset \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg2*)•Inf, sign(*fltArg2*)•pmax} if sign(*fltArg1*)=sign(*fltArg2*), {NaN, 0, -0} otherwise
fltArg1 infinity, *fltArg2* NaN, i.e.	$(-Inf, Inf)$	{NaN, 0, -0, *pmax*, *nmax*}

Table continues on the next page...

Table 3-102. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg3* (continued)

Subset of input domain	Worst-case error bounds for fltArg3 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg3 output
$\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = NaN \end{array} \right\}$		
fltArg1 infinity, *fltArg2* denormalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D \end{array} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg1*)•Inf, sign(*fltArg1*)•pmax}
fltArg1 infinity, *fltArg2* zero, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0 \end{array} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg1*)•Inf, sign(*fltArg1*)•pmax}
fltArg1 infinity, *fltArg2* normalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N \end{array} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg1*)•Inf, sign(*fltArg1*)•pmax}
fltArg1 NaN, *fltArg2* infinity, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = NaN, \\ fltIn_2 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
Both inputs NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = NaN, \\ fltIn_2 = NaN \end{array} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
fltArg1 NaN, *fltArg2* denormalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = NaN, \\ fltIn_2 \in D \end{array} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
fltArg1 NaN, *fltArg2* zero, i.e.	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}

Table continues on the next page...

Table 3-102. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg3* (continued)

Subset of input domain	Worst-case error bounds for fltArg3 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg3 output
$\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = NaN, \\ fltIn_2 \in D \end{array} \right\}$		
fltArg1 NaN, *fltArg2* normalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = NaN, \\ fltIn_2 \in N \end{array} \right\}$	$(-Inf, Inf)$	NaN
fltArg1 denormalized, *fltArg2* infinite, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg2*)•Inf, sign(*fltArg2*)•pmax}
fltArg1 denormalized, *fltArg2* NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 = NaN \end{array} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
Both inputs denormalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in D \end{array} \right\}$	$\langle -2^{23}, 2^{23} \rangle$	{0, -0}
fltArg1 zero, *fltArg2* infinite, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg2*)•Inf, sign(*fltArg2*)•pmax}
fltArg1 zero, *fltArg2* NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 = NaN \end{array} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
fltArg1 denormalized, *fltArg2* zero, i.e.	$\langle -0.5, 0.5 \rangle$	{0, -0}

Table continues on the next page...

Table 3-102. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg3* (continued)

Subset of input domain	Worst-case error bounds for fltArg3 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg3 output
$\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 = 0 \end{array} \right\}$		
fltArg1 zero, *fltArg2* denormalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in D \end{array} \right\}$	$\langle -1.5, 1.5 \rangle$	{0, -0}
fltArg1 normalized, *fltArg2* infinite, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{-sign(*fltArg2*)•Inf}
fltArg1 normalized, *fltArg2* NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 = NaN \end{array} \right\}$	$(-Inf, Inf)$	NaN
fltArg1 denormalized, *fltArg2* normalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N \end{array} \right\}$	$\langle -2^{23}, 2^{23} \rangle$	N/A
fltArg1 normalized, *fltArg2* denormalized, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N \end{array} \right\}$	$\langle -2^{23}, 2^{23} \rangle$	N/A
Both inputs normalized or zero, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \in N \cup \{-0, +0\} \end{array} \right\}$	$\langle -me_1, me_1 \rangle$	N/A
Both inputs normalized or zero, *fltArg3* output underflow i.e.	$(-Inf, Inf)$	{-0, +0}

Table continues on the next page...

Table 3-102. GMCLIB_ClarkInv_FLT Worst-Case Error Bounds - Output *fltArg3* (continued)

Subset of input domain	Worst-case error bounds for fltArg3 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg3 output
$\left\{ \begin{array}{c} (fltArg1,\ fltArg2) \in X \mid \\ X \in N \cup \{-0,\ +0\}, \\ \left(\begin{array}{c} \lvert ref\,fltArg2 \rvert - me_1 \cdot \mathrm{ULP}(ref\,fltArg2) \in D \wedge \\ ref\,fltArg2 \neq 0, \end{array} \right) \end{array} \right\}$		
Both inputs normalized or zero, intermediate result underflow i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2) \in X \mid \\ X \in N \cup \{-0,\ +0\}, \\ ir_1 \in D \vee \\ \left(\begin{array}{c} \lvert ir_2 \rvert - 1.5 \cdot \mathrm{ULP}(ir_2) \in D \wedge \\ ir_2 \neq 0, \end{array} \right) \end{array} \right\}$	$(-Inf, Inf)$	{NaN, -Inf, Inf}
Both inputs normalized or zero, *fltArg3* output overflow i.e. $\left\{ \begin{array}{c} (fltArg1,\ fltArg2) \in X \mid \\ X \in N \cup \{-0,\ +0\}, \\ \lvert ref\,fltArg2 \rvert + me_1 \cdot \mathrm{ULP}(ref\,fltArg2) \in M \end{array} \right\}$	$(-Inf, Inf)$	{NaN, -Inf, Inf}

3.33 Function GMCLIB_DecouplingPMSM_FLT

3.33.1 Declaration

```
void GMCLIB_DecouplingPMSM_FLT(SWLIBS_2Syst_FLT *const pUdqDec, const SWLIBS_2Syst_FLT *const
pUdq, const SWLIBS_2Syst_FLT *const pIdq, tFloat fltAngularVel, const
GMCLIB_DECOUPLINGPMSM_T_FLT *const pParam);
```

3.33.2 Arguments

Table 3-103. GMCLIB_DecouplingPMSM_FLT arguments

Type	Name	Direction	Description
SWLIBS_2Syst_FLT *const	pUdqDec	**output**	Pointer to the structure containing direct (u_{df_dec}) and quadrature (u_{qf_dec}) components of the decoupled stator voltage vector to be applied on the motor terminals.
const SWLIBS_2Syst_FLT *const	pUdq	**input**	Pointer to the structure containing direct (u_{df}) and quadrature (u_{qf}) components of the stator voltage vector generated by the current controllers.
const SWLIBS_2Syst_FLT *const	pIdq	**input**	Pointer to the structure containing direct (i_{df}) and quadrature (i_{qf}) components of the stator current vector measured on the motor terminals.
tFloat	fltAngularVel	**input**	Rotor angular velocity in rad/sec, referred to as ($omega_{ef}$) in the detailed section of the documentation.
const GMCLIB_DECOUPLINGPMSM_T_FLT *const	pParam	**input**	Pointer to the structure containing L_D and L_Q coefficients (see the detailed section of the documentation).

3.33.3 Worst-Case Error Bounds

Let $(UfltArg1, UfltArg2, IfltArg1, IfltArg2, fltAngularVel, fltLD, fltLQ) \in X$ be a vector of inputs to GMCLIB_DecouplingPMSM_FLT,

$reffltArg1$, $reffltArg2$ be the theoretical exact results for outputs $fltArg1$ and $fltArg2$, respectively,

$ira_1 = fltAngularVel \cdot IfltArg2$,

$ira_2 = ira_1 \cdot fltLQ$,

$irb_1 = fltAngularVel \cdot IfltArg1$,

$irb_2 = irb_1 \cdot fltLD$,

then

Table 3-104. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg1 Output

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
Any of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} is NaN, i.e.	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}

Table continues on the next page...

Table 3-104. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg1 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
$\left\{ \left(\begin{matrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{matrix} \right) \in X \middle\vert \begin{matrix} \\ \left\{ \begin{matrix} UfltArg1, IfltArg2, \\ fltLQ, \quad fltAngularVel \end{matrix} \right\} \cap NaN \neq \varnothing \end{matrix} \right\}$		
Any of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} is infinity, i.e. $\left\{ \left(\begin{matrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{matrix} \right) \in X \middle\vert \begin{matrix} \\ \left\{ \begin{matrix} UfltArg1, IfltArg2, \\ fltLQ, \quad fltAngularVel \end{matrix} \right\} \cap \{Inf, -Inf\} \neq \varnothing \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*, Inf, -Inf, 0, -0}
At least one of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} is denormalized, others are normalized, denormalized, or zero, i.e. $\left\{ \left(\begin{matrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{matrix} \right) \in X \middle\vert \begin{matrix} \\ \left\{ \begin{matrix} UfltArg1, IfltArg2, \\ fltLQ, \quad fltAngularVel \end{matrix} \right\} \cap D \neq \varnothing, \\ \left\{ \begin{matrix} UfltArg1, IfltArg2, \\ fltLQ, \quad fltAngularVel \end{matrix} \right\} \cap M = \varnothing \end{matrix} \right\}$	$(-Inf, Inf)$	{0, -0}
All of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} are zero, i.e. $\left\{ \left(\begin{matrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{matrix} \right) \in X \middle\vert \begin{matrix} \\ UfltArg1 = 0, \\ IfltArg2 = 0, \\ fltLQ = 0, \\ fltAngularVel = 0 \end{matrix} \right\}$	0	N/A
All of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} are normalized or zero, inputs lead to normalized or zero results, i.e.	$\langle -1, 1 \rangle$	N/A

Table continues on the next page...

Table 3-104. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg1 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
$\left\{ \begin{pmatrix} UfltArg1, UfltArg2,\\ IfltArg1, IfltArg2,\\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{\begin{array}{l} UfltArg1, IfltArg2,\\ fltLQ,\ \ fltAngularVel \end{array}\right\} \subset N \cup \{-0,\ +0\},\\ \{ira_1, ira_2\} \subset N \cup \{-0,\ +0\},\\ reffltArg1 = 0 \lor \\ \left(\begin{array}{c} 0.5 \cdot (\lvert reffltArg1 \rvert - \mathrm{ULP}(reffltArg1)) \notin D \land \\ \lvert reffltArg1 \rvert + \mathrm{ULP}(reffltArg1) \neq Inf \end{array}\right) \end{array} \right\}$		
All of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} are normalized or zero, *fltArg1* output overflow, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2,\\ IfltArg1, IfltArg2,\\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{\begin{array}{l} UfltArg1, IfltArg2,\\ fltLQ,\ \ fltAngularVel \end{array}\right\} \subset N \cup \{-0,\ +0\},\\ \lvert reffltArg1 \rvert + \mathrm{ULP}(reffltArg1) = Inf \end{array} \right\}$	$(-Inf, Inf)$	{sign(*reffltArg1*)•Inf, sign(*reffltArg1*)•pmax, NaN}
All of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} are normalized or zero, *fltArg1* output underflow or uncompensated calculation, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2,\\ IfltArg1, IfltArg2,\\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{\begin{array}{l} UfltArg1, IfltArg2,\\ fltLQ,\ \ fltAngularVel \end{array}\right\} \subset N \cup \{-0,\ +0\},\\ reffltArg1 \neq 0 \land \\ 0.5 \cdot (\lvert reffltArg1 \rvert - \mathrm{ULP}(reffltArg1)) \in D \end{array} \right\}$	$\langle -2^{24}, 2^{24} \rangle$	{0, -0}
All of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} are normalized or zero, intermediate result overflow, i.e.	$(-Inf, Inf)$	{*pmax*, *nmax*, Inf, -Inf}

Table continues on the next page...

Table 3-104. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg1 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
$\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{pmatrix} \in X \middle\vert \begin{matrix} \begin{Bmatrix} UfltArg1, IfltArg2, \\ fltLQ, \quad fltAngularVel \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ \begin{pmatrix} \lvert ira_1 \rvert + 0.5 \cdot \mathrm{ULP}(ira_1), \\ \lvert ira_2 \rvert + \mathrm{ULP}(ira_2) \end{pmatrix} \cap \{Inf, -Inf\} \neq \varnothing \end{matrix} \right\}$		
All of {*UfltArg1*, *IfltArg2*, *fltLQ*, *fltAngularVel*} are normalized or zero, intermediate result underflow, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{pmatrix} \in X \middle\vert \begin{matrix} \begin{Bmatrix} UfltArg1, IfltArg2, \\ fltLQ, \quad fltAngularVel \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ \begin{pmatrix} 0.5 \cdot (\lvert ira_1 \rvert - 0.5 \cdot \mathrm{ULP}(ira_1)) \in D \wedge \\ ira_1 \neq 0, \end{pmatrix} \vee \\ \begin{pmatrix} 0.5 \cdot (\lvert ira_2 \rvert - \mathrm{ULP}(ira_2)) \in D \wedge \\ ira_2 \neq 0, \end{pmatrix} \end{matrix} \right\}$	$\langle -2^{24}, 2^{24} \rangle$	{0, -0}

Table 3-105. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg2 Output

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
Any of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} is NaN, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{pmatrix} \in X \middle\vert \begin{Bmatrix} UfltArg2, IfltArg1, \\ fltLD, \quad fltAngularVel \end{Bmatrix} \cap NaN \neq \varnothing \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
Any of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} is infinity, i.e.	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*, Inf, -Inf, 0, -0}

Table continues on the next page...

Table 3-105. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{ \begin{array}{l} UfltArg2, IfltArg1, \\ fltLD,\quad fltAngularVel \end{array} \right\} \cap \{Inf, -Inf\} \neq \varnothing \end{array} \right\}$		
At least one of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} is denormalized, others are normalized, denormalized, or zero, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{ \begin{array}{l} UfltArg2, IfltArg1, \\ fltLD,\quad fltAngularVel \end{array} \right\} \cap D \neq \varnothing, \\ \left\{ \begin{array}{l} UfltArg2, IfltArg1, \\ fltLD,\quad fltAngularVel \end{array} \right\} \cap M = \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{0, -0}
All of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} are zero, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \;\middle|\; \begin{array}{c} UfltArg2 = 0, \\ IfltArg1 = 0, \\ fltLD = 0, \\ fltAngularVel = 0 \end{array} \right\}$	0	N/A
All of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} are normalized or zero, inputs lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{ \begin{array}{l} UfltArg2, IfltArg1, \\ fltLD,\quad fltAngularVel \end{array} \right\} \subset N \cup \{-0, +0\}, \\ \{irb_1, irb_2\} \subset N \cup \{-0, +0\}, \\ ref fltArg2 = 0 \vee \\ \begin{pmatrix} 0.5 \cdot (\lvert ref fltArg2 \rvert - \mathrm{ULP}(ref fltArg2)) \notin D \wedge \\ \lvert ref fltArg2 \rvert + \mathrm{ULP}(ref fltArg2) \neq Inf \end{pmatrix} \end{array} \right\}$	$\langle -1, 1 \rangle$	N/A

Table continues on the next page...

Table 3-105. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
All of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} are normalized or zero, *fltArg2* output overflow, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \,\middle\vert\, \begin{array}{c} \left\{ \begin{array}{l} UfltArg2, IfltArg1, \\ fltLD,\quad fltAngularVel \end{array} \right\} \subset N \cup \{-0,\ +0\}, \\ \lvert refltArg2 \rvert + \mathrm{ULP}(refltArg2) = Inf \end{array} \right\}$	$(-Inf, Inf)$	{sign(*reffltArg2*)•Inf, sign(*reffltArg2*)•pmax, NaN}
All of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} are normalized or zero, *fltArg2* output underflow or uncompensated calculation, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \,\middle\vert\, \begin{array}{c} \left\{ \begin{array}{l} UfltArg2, IfltArg1, \\ fltLD,\quad fltAngularVel \end{array} \right\} \subset N \cup \{-0,\ +0\}, \\ reffltArg2 \neq 0 \wedge \\ 0.5 \cdot (\lvert reffltArg2 \rvert - \mathrm{ULP}(reffltArg2)) \in D \end{array} \right\}$	$\langle -2^{24}, 2^{24} \rangle$	{0, -0}
All of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} are normalized or zero, intermediate result overflow, i.e. $\left\{ \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel,\ fltLD,\ fltLQ \end{pmatrix} \in X \,\middle\vert\, \begin{array}{c} \left\{ \begin{array}{l} UfltArg2, IfltArg1, \\ fltLD,\quad fltAngularVel \end{array} \right\} \subset N \cup \{-0,\ +0\}, \\ \begin{pmatrix} \lvert irb_1 \rvert + 0.5 \cdot \mathrm{ULP}(irb_1), \\ \lvert irb_2 \rvert + \mathrm{ULP}(irb_2) \end{pmatrix} \cap \{Inf,\ -Inf\} \neq \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{*pmax*, *nmax*, Inf, -Inf}
All of {*UfltArg2*, *IfltArg1*, *fltLD*, *fltAngularVel*} are normalized or zero, intermediate result underflow, i.e.	$\langle -2^{24}, 2^{24} \rangle$	{0, -0}

Table 3-105. GMCLIB_DecouplingPMSM_FLT Worst-Case Error Bounds - fltArg2 Output

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{ \begin{array}{c} \begin{pmatrix} UfltArg1, UfltArg2, \\ IfltArg1, IfltArg2, \\ fltAngularVel, fltLD, fltLQ \end{pmatrix} \in X \;\Big\vert \\ \left\{ \begin{array}{c} UfltArg2, IfltArg1, \\ fltLD, \quad fltAngularVel \end{array} \right\} \subset N \cup \{-0, +0\}, \\ \begin{pmatrix} 0.5 \cdot (\lvert irb_1 \rvert - 0.5 \cdot \mathrm{ULP}(irb_1)) \in D \wedge \\ ira_1 \neq 0, \end{pmatrix} \vee \\ \begin{pmatrix} 0.5 \cdot (\lvert irb_2 \rvert - \mathrm{ULP}(irb_2)) \in D \wedge \\ irb_2 \neq 0, \end{pmatrix} \end{array} \right)$		

3.34 Function GMCLIB_ElimDcBusRip_FLT

3.34.1 Declaration

```
void GMCLIB_ElimDcBusRip_FLT(SWLIBS_2Syst_FLT *const pOut, const SWLIBS_2Syst_FLT *const pIn,
const GMCLIB_ELIMDCBUSRIP_T_FLT *const pParam);
```

3.34.2 Arguments

Table 3-106. GMCLIB_ElimDcBusRip_FLT arguments

Type	Name	Direction	Description
SWLIBS_2Syst_FLT *const	pOut	**output**	Pointer to the structure with direct (α) and quadrature (β) components of the required stator voltage vector re-calculated so as to compensate for voltage ripples on the DC bus.
const SWLIBS_2Syst_FLT *const	pIn	**input**	Pointer to the structure with direct (α) and quadrature (β) components of the required stator voltage vector before compensation of voltage ripples on the DC bus.
const GMCLIB_ELIMDCBUSRIP_T_FLT *const	pParam	**input**	Pointer to the parameters structure.

3.34.3 Worst-Case Error Bounds

Let $(fltArg1, fltArg2, fltModIndex, fltArgDcBusMsr) \in X$ be a vector of inputs to GMCLIB_ElimDcBusRip_FLT,

reffltArg1, *reffltArg2* be the theoretical exact results for outputs *fltArg1* and *fltArg2*, respectively,

$ir_1 = fltArg1 \cdot fltModIndex$,

$ir_2 = fltArg2 \cdot fltModIndex$,

then

Table 3-107. GMCLIB_ElimDcBusRip_FLT Worst-Case Error Bounds - *fltArg1* Output

Subset of input domain	Worst-case error bounds for *fltArg1* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg1* output
Any of {*fltArg1*, *fltModIndex*, *fltArgDcBusMsr*} is NaN, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{Bmatrix} fltArg1, \\ fltModIndex, \\ fltArgDcBusMsr \end{Bmatrix} \cap NaN \neq \varnothing \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
Any of {*fltArg1*, *fltModIndex*, *fltArgDcBusMsr*} is infinity, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{Bmatrix} fltArg1, \\ fltModIndex, \\ fltArgDcBusMsr \end{Bmatrix} \cap \{Inf, -Inf\} \neq \varnothing \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*, Inf, -Inf}
At least one of {*fltArg1*, *fltModIndex*, *fltArgDcBusMsr*} is denormalized, others are normalized, denormalized, or zero, i.e.	$(-Inf, Inf)$	{*pmax*, *nmax*, Inf, -Inf, 0, -0}

Table continues on the next page...

Table 3-107. GMCLIB_ElimDcBusRip_FLT Worst-Case Error Bounds - *fltArg1* Output (continued)

Subset of input domain	Worst-case error bounds for *fltArg1* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg1* output
$\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{ \begin{array}{c} fltArg1, \\ fltModIndex, \\ fltArgDcBusMsr \end{array} \right\} \cap D \neq \varnothing, \\ \left\{ \begin{array}{c} fltArg1, \\ fltModIndex, \\ fltArgDcBusMsr \end{array} \right\} \cap M = \varnothing \end{array} \right\}$		
All of {*fltArg1*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, inputs lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{ \begin{array}{c} fltArg1, \\ fltModIndex \end{array} \right\} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ ir_1 \in N \cup \{-0, +0\}, \\ reffltArg1 = 0 \vee \\ \begin{pmatrix} \lvert reffltArg1 \rvert - 1.5 \cdot \mathrm{ULP}(reffltArg1) \notin D \wedge \\ \lvert reffltArg1 \rvert + 1.5 \cdot \mathrm{ULP}(reffltArg1) \neq Inf \end{pmatrix} \end{array} \right\}$	$\langle -1.5, 1.5 \rangle$	N/A
All of {*fltArg1*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is zero, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{array}{c} \left\{ \begin{array}{c} fltArg1, \\ fltModIndex \end{array} \right\} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr = 0 \end{array} \right\}$	0	{0, -0}
All of {*fltArg1*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, *fltArg1* output overflow, i.e.	$(-Inf, Inf)$	{sign(*reffltArg1*)•Inf, sign(*reffltArg1*)•pmax, NaN}

Table continues on the next page...

Table 3-107. GMCLIB_ElimDcBusRip_FLT Worst-Case Error Bounds - *fltArg1* Output (continued)

Subset of input domain	Worst-case error bounds for *fltArg1* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg1* output
$\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\vert \begin{matrix} \begin{Bmatrix} fltArg1, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ \lvert ref\,fltArg1 \rvert + 1.5 \cdot \mathrm{ULP}(ref\,fltArg1) = Inf \end{matrix} \right\}$		
All of {*fltArg1*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, *fltArg1* output underflow, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\vert \begin{matrix} \begin{Bmatrix} fltArg1, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ ref\,fltArg1 \neq 0 \wedge \\ \lvert ref\,fltArg1 \rvert - 1.5 \cdot \mathrm{ULP}(ref\,fltArg1) \in D \end{matrix} \right\}$	$\langle -1.5, 1.5 \rangle$	{0, -0}
All of {*fltArg1*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, intermediate result overflow, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\vert \begin{matrix} \begin{Bmatrix} fltArg1, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ \lvert ir_1 \rvert + 0.5 \cdot \mathrm{ULP}(ir_1) \in \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*, Inf, -Inf, 0, -0}
All of {*fltArg1*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, intermediate result underflow, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\vert \begin{matrix} \begin{Bmatrix} fltArg1, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ \lvert ir_1 \rvert - 0.5 \cdot \mathrm{ULP}(ir_1) \in D \wedge \\ ir_1 \neq 0 \end{matrix} \right\}$	$\langle -1.5, 1.5 \rangle$	{NaN, *pmax*, *nmax*, Inf, -Inf, 0, -0}

Table 3-108. GMCLIB_ElimDcBusRip_FLT Worst-Case Error Bounds - *fltArg2* Output

Subset of input domain	Worst-case error bounds for *fltArg2* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg2* output
Any of {*fltArg2*, *fltModIndex*, *fltArgDcBusMsr*} is NaN, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\| \begin{Bmatrix} fltArg2, \\ fltModIndex, \\ fltArgDcBusMsr \end{Bmatrix} \cap NaN \neq \varnothing \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*}
Any of {*fltArg2*, *fltModIndex*, *fltArgDcBusMsr*} is infinity, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\| \begin{Bmatrix} fltArg2, \\ fltModIndex, \\ fltArgDcBusMsr \end{Bmatrix} \cap \{Inf, -Inf\} \neq \varnothing \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*, Inf, -Inf}
At least one of {*fltArg2*, *fltModIndex*, *fltArgDcBusMsr*} is denormalized, others are normalized, denormalized, or zero, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\| \begin{array}{c} \begin{Bmatrix} fltArg2, \\ fltModIndex, \\ fltArgDcBusMsr \end{Bmatrix} \cap D \neq \varnothing, \\ \begin{Bmatrix} fltArg2, \\ fltModIndex, \\ fltArgDcBusMsr \end{Bmatrix} \cap M = \varnothing \end{array} \right\}$	$(-Inf, Inf)$	{*pmax*, *nmax*, Inf, -Inf, 0, -0}
All of {*fltArg2*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, inputs lead to normalized or zero results, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \middle\| \begin{array}{c} \begin{Bmatrix} fltArg2, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ ir_2 \in N \cup \{-0, +0\}, \\ reffltArg2 = 0 \vee \\ \begin{pmatrix} \|reffltArg2\| - 1.5 \cdot \mathrm{ULP}(reffltArg2) \notin D \wedge \\ \|reffltArg2\| + 1.5 \cdot \mathrm{ULP}(reffltArg2) \neq Inf \end{pmatrix} \end{array} \right\}$	$\langle -1.5, 1.5 \rangle$	N/A

Table continues on the next page...

Table 3-108. GMCLIB_ElimDcBusRip_FLT Worst-Case Error Bounds - *fltArg2* Output (continued)

Subset of input domain	Worst-case error bounds for *fltArg2* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg2* output
All of {*fltArg1*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is zero, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{matrix} \begin{Bmatrix} fltArg2, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr = 0 \end{matrix} \right\}$	0	{0, -0}
All of {*fltArg2*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, *fltArg2* output overflow, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{matrix} \begin{Bmatrix} fltArg2, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ \lvert reffltArg2 \rvert + 1.5 \cdot \mathrm{ULP}(reffltArg2) = Inf \end{matrix} \right\}$	$(-Inf, Inf)$	{sign(*reffltArg2*)•Inf, sign(*reffltArg2*)•pmax, NaN}
All of {*fltArg2*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, *fltArg2* output underflow, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{matrix} \begin{Bmatrix} fltArg2, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ reffltArg2 \neq 0 \land \\ \lvert reffltArg2 \rvert - 1.5 \cdot \mathrm{ULP}(reffltArg2) \in D \end{matrix} \right\}$	$\langle -1.5, 1.5 \rangle$	{0, -0}
All of {*fltArg2*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, intermediate result overflow, i.e. $\left\{ \begin{pmatrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{pmatrix} \in X \;\middle|\; \begin{matrix} \begin{Bmatrix} fltArg2, \\ fltModIndex \end{Bmatrix} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ \lvert ir_2 \rvert + 0.5 \cdot \mathrm{ULP}(ir_2) \in \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, *pmax*, *nmax*, Inf, -Inf, 0, -0}
All of {*fltArg2*, *fltModIndex*} are normalized or zero, *fltArgDcBusMsr* is normalized, intermediate result underflow, i.e.	$\langle -1.5, 1.5 \rangle$	{NaN, *pmax*, *nmax*, Inf, -Inf, 0, -0}

Table 3-108. GMCLIB_ElimDcBusRip_FLT Worst-Case Error Bounds - *fltArg2* Output

Subset of input domain	Worst-case error bounds for *fltArg2* output [ulp]	Allowed specific values (regardless the error bounds) for *fltArg2* output
$\left\{ \left(\begin{matrix} fltArg1, fltArg2, \\ fltModIndex, fltArgDcBusMsr \end{matrix} \right) \in X \middle\vert \begin{matrix} \left\{ \begin{matrix} fltArg2, \\ fltModIndex \end{matrix} \right\} \subset N \cup \{-0, +0\}, \\ fltArgDcBusMsr \in N, \\ \lvert ir_2 \rvert - 0.5 \cdot \mathrm{ULP}(ir_2) \in D \wedge \\ ir_2 \neq 0 \end{matrix} \right\}$		

3.35 Function GMCLIB_Park_FLT

3.35.1 Declaration

```
void GMCLIB_Park_FLT(SWLIBS_2Syst_FLT *pOut, const SWLIBS_2Syst_FLT *const pInAngle, const
SWLIBS_2Syst_FLT *const pIn);
```

3.35.2 Arguments

Table 3-109. GMCLIB_Park_FLT arguments

Type	Name	Direction	Description
SWLIBS_2Syst_FLT *	pOut	**input, output**	Pointer to the structure containing data of the two-phase rotational orthogonal system (d-q).
const SWLIBS_2Syst_FLT *const	pInAngle	**input**	Pointer to the structure where the values of the sine and cosine of the rotor position are stored.
const SWLIBS_2Syst_FLT *const	pIn	**input**	Pointer to the structure containing data of the two-phase stationary orthogonal system (α- β).

3.35.3 Worst-Case Error Bounds

Let $(AnglefltArg1, AnglefltArg2, InfltArg1, InfltArg2) \in X$ be a vector of inputs to GMCLIB_Park_FLT,

$reffltArg1$, $reffltArg2$ be the theoretical exact results for outputs $fltArg1$ and $fltArg2$, respectively,

$ir_1 = InfltArg1 \cdot AnglefltArg2$,

$ir_2 = InfltArg2 \cdot AnglefltArg1$,

$ir_3 = InfltArg2 \cdot AnglefltArg2$,

$ir_4 = -InfltArg1 \cdot AnglefltArg1$,

then

Table 3-110. GMCLIB_Park_FLT Worst-Case Error Bounds - fltArg1 Output

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle|\; X \cap M \neq \emptyset \right\}$	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*, -0, +0}
Any of the inputs is denormalized, neither is NaN or infinity, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle|\; X \cap D \neq \emptyset, X \cap M = \emptyset \right\}$	$(-Inf, Inf)$	N/A
All inputs normalized or zero, results normalized or zero, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle|\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2\} \subset N \cup \{-0, +0\}, \\ reffltArg1 = 0 \vee \\ \left(\begin{array}{c} \lvert reffltArg1 \rvert - \mathrm{ULP}(reffltArg1) \notin D \wedge \\ \lvert reffltArg1 \rvert + \mathrm{ULP}(reffltArg1) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -1, 1 \rangle$	N/A
All inputs zero, i.e.	0	N/A

Table continues on the next page...

Table 3-110. GMCLIB_Park_FLT Worst-Case Error Bounds - fltArg1 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
$\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\vert\; X \subset \{-0, +0\} \right\}$		
All inputs normalized or zero, *fltArg1* output overflow, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\vert\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \lvert reffltArg1 \rvert + \mathrm{ULP}(reffltArg1) = Inf \end{array} \right\}$	$\langle -1, 1 \rangle$	{sign(*reffltArg1*)•Inf, sign(*reffltArg1*)•pmax, NaN}
All inputs normalized or zero, *fltArg1* output underflow or underflow in internal compensation calculation, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\vert\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ 0.5 \cdot (\lvert reffltArg1 \rvert - \mathrm{ULP}(reffltArg1) \in D) \wedge \\ reffltArg1 \neq 0 \end{array} \right\}$	$\langle -2^{24}, 2^{24} \rangle$	{-0, +0}
All inputs normalized or zero, intermediate result overflow, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\vert\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ \lvert ir_1 \rvert \in \{Inf, -Inf\} \vee \lvert ir_2 \rvert \in \{Inf, -Inf\} \end{array} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
All inputs normalized or zero, intermediate result underflow, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\vert\; \begin{array}{c} X \subset N \cup \{-0, +0\}, \\ ir_1 \in D \vee ir_2 \in D \end{array} \right\}$	$(-Inf, Inf)$	N/A

Table 3-111. GMCLIB_Park_FLT Worst-Case Error Bounds - fltArg2 Output

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
Any of the inputs is NaN or infinity, i.e.	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*, -0, +0}

Table continues on the next page...

Table 3-111. GMCLIB_Park_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\| \begin{matrix} \\ X \cap M \neq \varnothing \end{matrix}\right\}$		
Any of the inputs is denormalized, neither is NaN or infinity, i.e. $\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\| \begin{matrix} \\ X \cap D \neq \varnothing, X \cap M = \varnothing \end{matrix}\right\}$	$(\text{-}Inf, Inf)$	N/A
All inputs normalized or zero, results normalized or zero, i.e. $\left\{\begin{matrix} \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \\ X \subset N \cup \{\text{-}0, +0\}, \\ \{ir_3, ir_4\} \subset N \cup \{\text{-}0, +0\}, \\ reffltArg2 = 0 \vee \\ \begin{pmatrix} \|reffltArg2\| - \mathrm{ULP}(reffltArg2) \notin D \wedge \\ \|reffltArg2\| + \mathrm{ULP}(reffltArg2) \neq Inf \end{pmatrix} \end{matrix}\right\}$	$\langle\text{-}1, 1\rangle$	N/A
All inputs zero, i.e. $\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\| \begin{matrix} \\ X \subset \{\text{-}0, +0\} \end{matrix}\right\}$	0	N/A
All inputs normalized or zero, *fltArg2* output overflow, i.e. $\left\{\begin{matrix} \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \\ X \subset N \cup \{\text{-}0, +0\}, \\ \|reffltArg2\| + \mathrm{ULP}(reffltArg2) = Inf \end{matrix}\right\}$	$\langle\text{-}1, 1\rangle$	{sign(*reffltArg2*)•Inf, sign(*reffltArg2*)•pmax, NaN}
All inputs normalized or zero, *fltArg2* output underflow or underflow in internal compensation calculation, i.e. $\left\{\begin{matrix} \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \\ X \subset N \cup \{\text{-}0, +0\}, \\ 0.5 \cdot (\|reffltArg2\| - \mathrm{ULP}(reffltArg2) \in D) \wedge \\ reffltArg2 \neq 0 \end{matrix}\right\}$	$\langle\text{-}2^{24}, 2^{24}\rangle$	{-0, +0}
All inputs normalized or zero, intermediate result overflow, i.e.	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

Table continues on the next page...

Table 3-111. GMCLIB_Park_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert ir_3 \rvert \in \{Inf, -Inf\} \vee \lvert ir_4 \rvert \in \{Inf, -Inf\} \end{matrix} \right\}$		
All inputs normalized or zero, intermediate result underflow, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ ir_3 \in D \vee ir_4 \in D \end{matrix} \right\}$	$(-Inf, Inf)$	N/A

3.36 Function GMCLIB_ParkInv_FLT

3.36.1 Declaration

```
void GMCLIB_ParkInv_FLT(SWLIBS_2Syst_FLT *const pOut, const SWLIBS_2Syst_FLT *const pInAngle,
const SWLIBS_2Syst_FLT *const pIn);
```

3.36.2 Arguments

Table 3-112. GMCLIB_ParkInv_FLT arguments

Type	Name	Direction	Description
SWLIBS_2Syst_FLT *const	pOut	**input, output**	Pointer to the structure containing data of the two-phase stationary orthogonal system (α- β).
const SWLIBS_2Syst_FLT *const	pInAngle	**input**	Pointer to the structure where the values of the sine and cosine of the rotor position are stored.
const SWLIBS_2Syst_FLT *const	pIn	**input**	Pointer to the structure containing data of the two-phase rotational orthogonal system (d-q).

3.36.3 Worst-Case Error Bounds

Let (*AnglefltArg1*, *AnglefltArg2*, *InfltArg1*, *InfltArg2*) ∈ *X* be a vector of inputs to GMCLIB_ParkInv_FLT,

reffltArg1, *reffltArg2* be the theoretical exact results for outputs *fltArg1* and *fltArg2*, respectively,

$ir_1 = InfltArg1 \cdot AnglefltArg2$,

$ir_2 = -InfltArg2 \cdot AnglefltArg1$,

$ir_3 = InfltArg1 \cdot AnglefltArg1$,

$ir_4 = InfltArg2 \cdot AnglefltArg2$,

then

Table 3-113. GMCLIB_ParkInv_FLT Worst-Case Error Bounds - fltArg1 Output

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
Any of the inputs is NaN or infinity, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle|\; X \cap M \neq \varnothing \right\}$	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*, -0, +0}
Any of the inputs is denormalized, neither is NaN or infinity, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle|\; X \cap D \neq \varnothing, X \cap M = \varnothing \right\}$	$(-Inf, Inf)$	N/A
All inputs normalized or zero, results normalized or zero, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \{ir_1, ir_2\} \subset N \cup \{-0, +0\}, \\ reffltArg1 = 0 \vee \\ \begin{pmatrix} \lvert reffltArg1 \rvert - \mathrm{ULP}(reffltArg1) \notin D \wedge \\ \lvert reffltArg1 \rvert + \mathrm{ULP}(reffltArg1) \neq Inf \end{pmatrix} \end{matrix} \right\}$	$\langle -1, 1 \rangle$	N/A
All inputs zero, i.e.	0	N/A

Table continues on the next page...

Table 3-113. GMCLIB_ParkInv_FLT Worst-Case Error Bounds - fltArg1 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg1 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg1 output
$\left\{ \begin{pmatrix} AngleFltArg1, AngleFltArg2, \\ InFltArg1, InFltArg2 \end{pmatrix} \in X \;\middle|\; X \subset \{-0, +0\} \right\}$		
All inputs normalized or zero, *fltArg1* output overflow, i.e. $\left\{ \begin{pmatrix} AngleFltArg1, AngleFltArg2, \\ InFltArg1, InFltArg2 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert refFltArg1 \rvert + \mathrm{ULP}(refFltArg1) = Inf \end{matrix} \right\}$	$\langle -1, 1 \rangle$	{sign(*reffltArg1*)•Inf, sign(*reffltArg1*)•pmax, NaN}
All inputs normalized or zero, *fltArg1* output underflow or underflow in internal compensation calculation, i.e. $\left\{ \begin{pmatrix} AngleFltArg1, AngleFltArg2, \\ InFltArg1, InFltArg2 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ 0.5 \cdot (\lvert refFltArg1 \rvert - \mathrm{ULP}(refFltArg1) \in D) \wedge \\ refFltArg1 \neq 0 \end{matrix} \right\}$	$\langle -2^{24}, 2^{24} \rangle$	{-0, +0}
All inputs normalized or zero, intermediate result overflow, i.e. $\left\{ \begin{pmatrix} AngleFltArg1, AngleFltArg2, \\ InFltArg1, InFltArg2 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert ir_1 \rvert \in \{Inf, -Inf\} \vee \lvert ir_2 \rvert \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)
All inputs normalized or zero, intermediate result underflow, i.e. $\left\{ \begin{pmatrix} AngleFltArg1, AngleFltArg2, \\ InFltArg1, InFltArg2 \end{pmatrix} \in X \;\middle|\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ ir_1 \in D \vee ir_2 \in D \end{matrix} \right\}$	$(-Inf, Inf)$	N/A

Table 3-114. GMCLIB_ParkInv_FLT Worst-Case Error Bounds - fltArg2 Output

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
Any of the inputs is NaN or infinity, i.e.	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*, -0, +0}

Table continues on the next page...

Table 3-114. GMCLIB_ParkInv_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\vert \begin{matrix} \\ X \cap M \neq \varnothing \end{matrix}\right\}$		
Any of the inputs is denormalized, neither is NaN or infinity, i.e. $\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\vert \begin{matrix} \\ X \cap D \neq \varnothing, X \cap M = \varnothing \end{matrix}\right\}$	$(-Inf, Inf)$	N/A
All inputs normalized or zero, results normalized or zero, i.e. $\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\vert \begin{matrix} \\ X \subset N \cup \{-0, +0\}, \\ \{ir_3, ir_4\} \subset N \cup \{-0, +0\}, \\ reffltArg2 = 0 \vee \\ \begin{pmatrix} \lvert reffltArg2\rvert - \mathrm{ULP}(reffltArg2) \notin D \wedge \\ \lvert reffltArg2\rvert + \mathrm{ULP}(reffltArg2) \neq Inf \end{pmatrix} \end{matrix}\right\}$	$\langle -1, 1\rangle$	N/A
All inputs zero, i.e. $\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\vert \begin{matrix} \\ X \subset \{-0, +0\} \end{matrix}\right\}$	0	N/A
All inputs normalized or zero, *fltArg2* output overflow, i.e. $\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\vert \begin{matrix} \\ X \subset N \cup \{-0, +0\}, \\ \lvert reffltArg2\rvert + \mathrm{ULP}(reffltArg2) = Inf \end{matrix}\right\}$	$\langle -1, 1\rangle$	{sign(*reffltArg2*)•Inf, sign(*reffltArg2*)•pmax, NaN}
All inputs normalized or zero, *fltArg2* output underflow or underflow in internal compensation calculation, i.e. $\left\{\begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \middle\vert \begin{matrix} \\ X \subset N \cup \{-0, +0\}, \\ 0.5 \cdot (\lvert reffltArg2\rvert - \mathrm{ULP}(reffltArg2) \in D) \wedge \\ reffltArg2 \neq 0 \end{matrix}\right\}$	$\langle -2^{24}, 2^{24}\rangle$	{-0, +0}
All inputs normalized or zero, intermediate result overflow, i.e.	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

Table continues on the next page...

Table 3-114. GMCLIB_ParkInv_FLT Worst-Case Error Bounds - fltArg2 Output (continued)

Subset of input domain	Worst-case error bounds for fltArg2 output [ulp]	Allowed specific values (regardless the error bounds) for fltArg2 output
$\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\vert\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ \lvert ir_3 \rvert \in \{Inf, -Inf\} \vee \lvert ir_4 \rvert \in \{Inf, -Inf\} \end{matrix} \right\}$		
All inputs normalized or zero, intermediate result underflow, i.e. $\left\{ \begin{pmatrix} AnglefltArg1, AnglefltArg2, \\ InfltArg1, InfltArg2 \end{pmatrix} \in X \;\middle\vert\; \begin{matrix} X \subset N \cup \{-0, +0\}, \\ ir_3 \in D \vee ir_4 \in D \end{matrix} \right\}$	$(-Inf, Inf)$	N/A

3.37 Function GMCLIB_SvmStd_FLT

3.37.1 Declaration

```
tU32 GMCLIB_SvmStd_FLT(SWLIBS_3Syst_FLT *pOut, const SWLIBS_2Syst_FLT *const pIn);
```

3.37.2 Arguments

Table 3-115. GMCLIB_SvmStd_FLT arguments

Type	Name	Direction	Description
SWLIBS_3Syst_FLT *	pOut	**input, output**	Pointer to the structure containing calculated duty-cycle ratios of the 3-Phase system.
const SWLIBS_2Syst_FLT *const	pIn	**input**	Pointer to the structure containing direct U_{α} and quadrature U_{β} components of the stator voltage vector.

3.37.3 Worst-Case Error Bounds

Note: The allowed error bounds for the floating-point outputs of this function are expressed in terms of allowed absolute error.

Let $(fltArg1, fltArg2) \in X$ be a vector of inputs to GMCLIB_SvmStd_FLT,

then

Table 3-116. GMCLIB_SvmStd_FLT Return Value

Subset of input domain	Meaningful Return Value
Any of the inputs is NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltArg1 = NaN \vee \\ fltArg2 = NaN \end{array} \right\}$	No
Normalized, denormalized, or zero inputs, at least one input is outside the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert > 1 \vee \lvert fltArg2 \rvert > 1 \end{array} \right\}$	No
All inputs within the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert \le 1, \lvert fltArg2 \rvert \le 1 \end{array} \right\}$	Yes

Table 3-117. GMCLIB_SvmStd_FLT Worst-Case Error Bounds - Output fltArg1

Subset of input domain	Worst-case absolute error for fltArg1 output	Allowed specific values (regardless the error bounds) for fltArg1 output
Any of the inputs is NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltArg1 = NaN \vee \\ fltArg2 = NaN \end{array} \right\}$	Inf	NaN
Normalized, denormalized, or zero inputs, at least one input is outside the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert > 1 \vee \lvert fltArg2 \rvert > 1 \end{array} \right\}$	Inf	{NaN, Inf, -Inf}

Table continues on the next page...

Table 3-117. GMCLIB_SvmStd_FLT Worst-Case Error Bounds - Output fltArg1 (continued)

Subset of input domain	Worst-case absolute error for fltArg1 output	Allowed specific values (regardless the error bounds) for fltArg1 output
All inputs within the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert \le 1, \lvert fltArg2 \rvert \le 1 \end{array} \right\}$	2^{-19}	N/A

Table 3-118. GMCLIB_SvmStd_FLT Worst-Case Error Bounds - Output fltArg2

Subset of input domain	Worst-case absolute error for fltArg2 output	Allowed specific values (regardless the error bounds) for fltArg2 output
Any of the inputs is NaN, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltArg1 = NaN \vee \\ fltArg2 = NaN \end{array} \right\}$	Inf	NaN
Normalized, denormalized, or zero inputs, at least one input is outside the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert > 1 \vee \lvert fltArg2 \rvert > 1 \end{array} \right\}$	Inf	{NaN, Inf, -Inf}
All inputs within the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert \le 1, \lvert fltArg2 \rvert \le 1 \end{array} \right\}$	2^{-19}	N/A

Table 3-119. GMCLIB_SvmStd_FLT Worst-Case Error Bounds - Output fltArg3

Subset of input domain	Worst-case absolute error for fltArg3 output	Allowed specific values (regardless the error bounds) for fltArg3 output
Any of the inputs is NaN, i.e.	Inf	NaN

Table continues on the next page...

Table 3-119. GMCLIB_SvmStd_FLT Worst-Case Error Bounds - Output fltArg3 (continued)

Subset of input domain	Worst-case absolute error for fltArg3 output	Allowed specific values (regardless the error bounds) for fltArg3 output
$\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ fltArg1 = NaN \vee \\ fltArg2 = NaN \end{array} \right\}$		
Normalized, denormalized, or zero inputs, at least one input is outside the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert > 1 \vee \lvert fltArg2 \rvert > 1 \end{array} \right\}$	Inf	{NaN, Inf, -Inf}
All inputs within the [-1, 1] interval, i.e. $\left\{ \begin{array}{c} (fltArg1, fltArg2) \in X \mid \\ X \subset N \cup D \cup \{-0, +0\}, \\ \lvert fltArg1 \rvert \le 1, \lvert fltArg2 \rvert \le 1 \end{array} \right\}$	2^{-19}	N/A

3.38 Function MLIB_Abs_FLT

3.38.1 Declaration

```
INLINEtFloat MLIB_Abs_FLT(register tFloat fltIn);
```

3.38.2 Arguments

Table 3-120. MLIB_Abs_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn	**input**	Input value.

3.38.3 Worst-Case Error Bounds

Let $fltIn \in X$ be an input to MLIB_Abs_FLT,

then

Table 3-121. MLIB_Abs_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input infinity, i.e. $\{fltIn \in X \mid fltIn \in \{Inf, -Inf\}\}$	N/A	{Inf, *pmax*}
Input NaN, i.e. $\{fltIn \in X \mid fltIn = NaN\}$	N/A	{NaN, *pmax*}
Denormalized input, i.e. $\{fltIn \in X \mid X \subset D\}$	0	0
Normalized input, i.e. $\{fltIn \in X \mid X \subset N\}$	0	N/A
Zero input, i.e. $\{fltIn \in X \mid fltIn = 0\}$	0	{-0, +0}

3.39 Function MLIB_Add_FLT

3.39.1 Declaration

```
INLINEtFloat MLIB_Add_FLT(register tFloat fltIn1, register tFloat fltIn2);
```

3.39.2 Arguments

Table 3-122. MLIB_Add_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	First value to be add.
register tFloat	fltIn2	**input**	Second value to be add.

3.39.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2) \in X$ be a vector of inputs to MLIB_Add_FLT,

refResult be the theoretical exact result,

then

Table 3-123. MLIB_Add_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Both inputs infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*}
fltIn1 arbitrary value, *fltIn2* NaN, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_2 = NaN \}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 infinity, *fltIn2* normalized, denormalized, or zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N \cup D \cup \{-0, +0\} \end{matrix} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 NaN, *fltIn2* arbitrary value, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN \}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 normalized, denormalized, or zero, *fltIn2* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N \cup D \cup \{-0, +0\}, \\ fltIn_2 \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	{sign(*fltIn2*)•Inf, sign(*fltIn2*)•*pmax*}
Both inputs denormalized, i.e. $\{ fltIn_1, fltIn_2 \in X \mid X \subset D \}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* zero, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, fltIn_2 = 0 \}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* normalized, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in N \end{matrix} \right\}$	$\langle -(2^{24} - 1), 2^{24} - 1 \rangle$	N/A
fltIn1 zero, *fltIn2* denormalized, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 \in D \}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
Both inputs zero, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 = 0 \}$	0	N/A
fltIn1 zero, *fltIn2* normalized, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in N \end{matrix} \right\}$	0	N/A
fltIn1 normalized, *fltIn2* denormalized, i.e.		N/A

Table continues on the next page...

Table 3-123. MLIB_Add_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N, fltIn_2 \in D \end{array} \right\}$	$\langle -(2^{24} - 1), 2^{24} - 1 \rangle$	
fltIn1 normalized, *fltIn2* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N, fltIn_2 = 0 \end{array} \right\}$	0	N/A
Both inputs normalized, result normalized or zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
Both inputs normalized, result overflow, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) = Inf \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{sign(*refResult*)•Inf, sign(*refResult*)•*pmax*}
Both inputs normalized, result underflow, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}

3.40 Function MLIB_Convert_F32FLT

3.40.1 Declaration

```
INLINEtFrac32 MLIB_Convert_F32FLT(register tFloat fltIn1, register tFloat fltIn2);
```

3.40.2 Arguments

Table 3-124. MLIB_Convert_F32FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Input value in single precision floating point format to be converted.
register tFloat	fltIn2	**input**	Scale factor in single precision floating point format.

3.40.3 Worst-Case Error Bounds

This function returns a fixed-point result as specified in the Automotive Math and Motor Control Library Set for NXP S32K14x devices User's Guide. If any of the inputs is a NaN or infinity, then the function returns an arbitrary value.

3.41 Function MLIB_Convert_F16FLT

3.41.1 Declaration

```
INLINEtFrac16 MLIB_Convert_F16FLT(register tFloat fltIn1, register tFloat fltIn2);
```

3.41.2 Arguments

Table 3-125. MLIB_Convert_F16FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Input value in single precision floating point format to be converted.
register tFloat	fltIn2	**input**	Scale factor in single precision floating point format.

3.41.3 Worst-Case Error Bounds

This function returns a fixed-point result as specified in the Automotive Math and Motor Control Library Set for NXP S32K14x devices User's Guide. If any of the inputs is a NaN or infinity, then the function returns an arbitrary value.

3.42 Function MLIB_Convert_FLTF16

3.42.1 Declaration

```
INLINEtFloat MLIB_Convert_FLTF16(register tFrac16 f16In1, register tFrac16 f16In2);
```

3.42.2 Arguments

Table 3-126. MLIB_Convert_FLTF16 arguments

Type	Name	Direction	Description
register tFrac16	f16In1	**input**	Input value in 16-bit fractional format to be converted.
register tFrac16	f16In2	**input**	Scale factor in 16-bit fractional format.

3.42.3 Worst-Case Error Bounds

This function returns a result with worst-case error bounds of $\langle -1536, 1536\rangle$ ulp. The result is always a normalized value or zero.

3.43 Function MLIB_Convert_FLTF32

3.43.1 Declaration

```
INLINEtFloat MLIB_Convert_FLTF32(register tFrac32 f32In1, register tFrac32 f32In2);
```

3.43.2 Arguments

Table 3-127. MLIB_Convert_FLTF32 arguments

Type	Name	Direction	Description
register tFrac32	f32In1	**input**	Input value in 32-bit fractional format to be converted.
register tFrac32	f32In2	**input**	Scale factor in 32-bit fractional format.

3.43.3 Worst-Case Error Bounds

This function returns a result with worst-case error bounds of $\langle -1536, 1536\rangle$ ulp. The result is always a normalized value or zero.

3.44 Function MLIB_ConvertPU_F32FLT

3.44.1 Declaration

```
INLINEtFrac32 MLIB_ConvertPU_F32FLT(register tFloat fltIn);
```

3.44.2 Arguments

Table 3-128. MLIB_ConvertPU_F32FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn	**input**	Input value in single precision floating point format to be converted.

3.44.3 Worst-Case Error Bounds

This function returns a fixed-point result as specified in the Automotive Math and Motor Control Library Set for NXP S32K14x devices User's Guide. If any of the inputs is a NaN or infinity, then the function returns an arbitrary value.

3.45 Function MLIB_ConvertPU_F16FLT

3.45.1 Declaration

```
INLINEtFrac16 MLIB_ConvertPU_F16FLT(register tFloat fltIn);
```

3.45.2 Arguments

Table 3-129. MLIB_ConvertPU_F16FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn	**input**	Input value in single precision floating point format to be converted.

3.45.3 Worst-Case Error Bounds

This function returns a fixed-point result as specified in the Automotive Math and Motor Control Library Set for NXP S32K14x devices User's Guide. If any of the inputs is a NaN or infinity, then the function returns an arbitrary value.

3.46 Function MLIB_ConvertPU_FLTF16

3.46.1 Declaration

```
INLINEtFloat MLIB_ConvertPU_FLTF16(register tFrac16 f16In);
```

3.46.2 Arguments

Table 3-130. MLIB_ConvertPU_FLTF16 arguments

Type	Name	Direction	Description
register tFrac16	f16In	**input**	Input value in 16-bit fractional format to be converted.

3.46.3 Worst-Case Error Bounds

This function returns a result with worst-case error bounds of $\langle -1536, 1536 \rangle$ ulp. The result is always a normalized value or zero.

3.47 Function MLIB_ConvertPU_FLTF32

3.47.1 Declaration

```
INLINEtFloat MLIB_ConvertPU_FLTF32(register tFrac32 f32In);
```

3.47.2 Arguments

Table 3-131. MLIB_ConvertPU_FLTF32 arguments

Type	Name	Direction	Description
register tFrac32	f32In	**input**	Input value in 32-bit fractional format to be converted.

3.47.3 Worst-Case Error Bounds

This function returns a result with worst-case error bounds of $\langle -1536, 1536 \rangle$ ulp. The result is always a normalized value or zero.

3.48 Function MLIB_Div_FLT

3.48.1 Declaration

```
INLINEtFloat MLIB_Div_FLT(register tFloat fltIn1, register tFloat fltIn2);
```

3.48.2 Arguments

Table 3-132. MLIB_Div_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Numerator of division.
register tFloat	fltIn2	**input**	Denominator of division.

3.48.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2) \in X$ be a vector of inputs to MLIB_Div_FLT,

refResult be the theoretical exact result,

then

Table 3-133. MLIB_Div_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Both inputs infinity, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in \{Inf, -Inf\} \right\}$	N/A	{NaN, 0, -0}
fltIn1 infinity, *fltIn2* NaN, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 = NaN \right\}$	N/A	{NaN, 0, -0}
fltIn1 infinity, *fltIn2* NaN, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in D \right\}$	N/A	{Inf, *pmax*} if sign(*fltIn1*) = sign(*fltIn2*), {-Inf, *nmax*} if sign(*fltIn1*) ≠ sign(*fltIn2*)
fltIn1 infinity, *fltIn2* zero, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 = 0 \right\}$	N/A	{Inf, *pmax*} if sign(*fltIn1*) = sign(*fltIn2*), {-Inf, *nmax*} if sign(*fltIn1*) ≠ sign(*fltIn2*)
fltIn1 infinity, *fltIn2* normalized, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in N \right\}$	N/A	{Inf, *pmax*} if sign(*fltIn1*) = sign(*fltIn2*), {-Inf, *nmax*} if sign(*fltIn1*) ≠ sign(*fltIn2*)
fltIn1 NaN, *fltIn2* infinity, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN, fltIn_2 \in \{Inf, -Inf\} \right\}$	N/A	{NaN, 0, -0}
Both inputs NaN, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN, fltIn_2 = NaN \right\}$	N/A	{NaN, 0, -0}
fltIn1 NaN, *fltIn2* denormalized, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN, fltIn_2 \in D \right\}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 NaN, *fltIn2* zero, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN, fltIn_2 = 0 \right\}$	N/A	{NaN, *pmax*, *nmax*}

Table continues on the next page...

Table 3-133. MLIB_Div_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 NaN, *fltIn2* normalized, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN, fltIn_2 \in N\}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 denormalized, *fltIn2* infinity, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in \{Inf, -Inf\} \end{matrix}\right\}$	0	N/A
fltIn1 denormalized, *fltIn2* NaN, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, \\ fltIn_2 = NaN \end{matrix}\right\}$	N/A	{NaN, 0, -0}
Both inputs denormalized, i.e. $\{fltIn_1, fltIn_2 \in X \mid X \subset D\}$	$\langle -0.5, 0.5 \rangle$	NaN or (*pmax* if sign(*fltIn1*) = sign(*fltIn2*), *nmax* if sign(*fltIn1*) ≠ sign(*fltIn2*))
fltIn1 denormalized, *fltIn2* zero, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, fltIn_2 = 0\}$	N/A	{Inf, *pmax*} if sign(*fltIn1*) = sign(*fltIn2*), {-Inf, *nmax*} if sign(*fltIn1*) ≠ sign(*fltIn2*)
fltIn1 denormalized, *fltIn2* normalized, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in N \end{matrix}\right\}$	$\langle -0.5, 0.5 \rangle$	+0 if sign(*fltIn1*) = sign(*fltIn2*), -0 if sign(*fltIn1*) ≠ sign(*fltIn2*)
fltIn1 zero, *fltIn2* infinity, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\} \end{matrix}\right\}$	0	N/A
fltIn1 zero, *fltIn2* NaN, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 = NaN\}$	N/A	{NaN, 0, -0}
fltIn1 zero, *fltIn2* denormalized, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 \in D\}$	0	*pmax* if sign(*fltIn1*) = sign(*fltIn2*), *nmax* if sign(*fltIn1*) ≠ sign(*fltIn2*)
Both inputs zero, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 = 0\}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* normalized, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 \in N\}$	0	N/A
fltIn1 normalized, *fltIn2* infinity, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\} \end{matrix}\right\}$	0	N/A
fltIn1 normalized, *fltIn2* NaN, i.e.	N/A	{NaN, 0, -0}

Table continues on the next page...

Table 3-133. MLIB_Div_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in N, fltIn_2 = NaN \}$		
fltIn1 normalized, *fltIn2* denormalized, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in N, fltIn_2 \in D \}$	$\langle -0.5, 0.5 \rangle$	{Inf, *pmax*} if sign(*fltIn1*) = sign(*fltIn2*), {-Inf, *nmax*} if sign(*fltIn1*) ≠ sign(*fltIn2*)
fltIn1 normalized, *fltIn2* zero, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in N, fltIn_2 = 0 \}$	N/A	{Inf, *pmax*} if sign(*fltIn1*) = sign(*fltIn2*), {-Inf, *nmax*} if sign(*fltIn1*) ≠ sign(*fltIn2*)
Both inputs normalized, normalized result, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \left(\begin{matrix} \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{matrix} \right) \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
Both inputs normalized, result overflow, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) = Inf \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	{Inf, *pmax*} if sign(*fltIn1*) = sign(*fltIn2*), {-Inf, *nmax*} if sign(*fltIn1*) ≠ sign(*fltIn2*)
Both inputs normalized, result underflow, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	+0 if sign(*fltIn1*) = sign(*fltIn2*), -0 if sign(*fltIn1*) ≠ sign(*fltIn2*)

3.49 Function MLIB_Mac_FLT

3.49.1 Declaration

```
INLINEtFloat MLIB_Mac_FLT(register tFloat fltIn1, register tFloat fltIn2, register tFloat fltIn3);
```

3.49.2 Arguments

Table 3-134. MLIB_Mac_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Input value to be add.

Table continues on the next page...

Table 3-134. MLIB_Mac_FLT arguments (continued)

Type	Name	Direction	Description
register tFloat	fltIn2	**input**	First value to be multiplied.
register tFloat	fltIn3	**input**	Second value to be multiplied.

3.49.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2, fltIn3) \in X$ be a vector of inputs to MLIB_Mac_FLT,

$mr = fltIn_2 \cdot fltIn_3$,

refResult be the theoretical exact result,

then

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs infinity, i.e. $\{fltIn_1, fltIn_2, fltIn_3 \in X \mid X \subset \{Inf, -Inf\}\}$	N/A	{Inf, *pmax*} if (sign(*fltIn2*)=sign(*fltIn3*) and *fltIn1* > 0) or (sign(*fltIn2*)≠sign(*fltIn3*) and *fltIn1* < 0), {-Inf, *nmax*} otherwise
fltIn1 zero or denormalized, *fltIn2* and *fltIn3* infinity, i.e. $\left\{\begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D \bigvee fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in \{Inf, -Inf\} \end{array}\right\}$	N/A	{Inf, *pmax*} if sign(*fltIn2*)=sign(*fltIn3*), {-Inf, *nmax*} otherwise
fltIn1 normalized, *fltIn2* and *fltIn3* infinity, i.e. $\left\{\begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in \{Inf, -Inf\} \end{array}\right\}$	$(-Inf, Inf)$	Inf if sign(*fltIn2*)=sign(*fltIn3*), -Inf otherwise
fltIn1 NaN, i.e. $\{fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = NaN\}$	$(-Inf, Inf)$	NaN
fltIn2 NaN, i.e. $\{fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_2 = NaN\}$	$(-Inf, Inf)$	NaN
fltIn3 NaN, i.e. $\{fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_3 = NaN\}$	$(-Inf, Inf)$	NaN

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn2 zero, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_2 = 0, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0}
fltIn2 infinity, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 = 0 \end{array} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0}
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, fltIn_2 \in D, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, fltIn_2 \in D, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, Inf, -Inf}
fltIn1 infinity, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, Inf, -Inf}
fltIn1 zero, *fltIn2* normalized, *fltIn3* infinity, i.e.	N/A	{Inf, *pmax*} if sign(*fltIn2*)=sign(*fltIn3*), {-Inf, *nmax*} otherwise

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$		
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{Inf, *pmax*} if sign(*fltIn2*)=sign(*fltIn3*), {-Inf, *nmax*} otherwise
fltIn1 normalized, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	Inf if sign(*fltIn2*)=sign(*fltIn3*), -Inf otherwise
fltIn1 infinity, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
All inputs zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 denormalized, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$	0	N/A

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{array} \right\}$	0	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{array} \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* zero, i.e.	0	{-0, +0}

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \middle\vert \begin{matrix} fltIn_1 \in D, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{matrix} \right\}$		
fltIn1 normalized, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \middle\vert \begin{matrix} fltIn_1 \in N, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{matrix} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \middle\vert \begin{matrix} fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \middle\vert \begin{matrix} fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 denormalized, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \middle\vert \begin{matrix} fltIn_1 \in D, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 normalized, *fltIn2* infinity, *fltIn3* denormalized, i.e.	$(-Inf, Inf)$	{NaN, Inf, -Inf, 0, -0}

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{array} \right\}$		
fltIn1 infinity, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = \in D, fltIn_3 \in D \end{array} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in D, fltIn_3 \in D \end{array} \right\}$	0	N/A
All inputs denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid X \subset D, \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* and *fltIn3* denormalized, i.e.	0	N/A

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, fltIn_2 \in D, fltIn_3 \in D \right\}$		
fltIn1 infinity, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in N, fltIn_3 \in D \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, fltIn_2 \in N, fltIn_3 \in D \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, fltIn_2 \in N, fltIn_3 \in D \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 normalized, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, fltIn_2 \in N, fltIn_3 \in D \right\}$	$(-Inf, Inf)$	N/A
fltIn1 infinity, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	N/A	{NaN, Inf, -Inf, 0, -0, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	N/A	{Inf, -Inf, *pmax*, *nmax*}
fltIn1 denormalized, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	N/A	{Inf, -Inf, *pmax*, *nmax*}
fltIn1 normalized, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	$(-Inf, Inf)$	{Inf, -Inf}

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 infinity, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in \{Inf, -Inf\}, fltIn_2 = 0, fltIn_3 \in N \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 = 0, fltIn_2 = 0, fltIn_3 \in N \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in D, fltIn_2 = 0, fltIn_3 \in N \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 = 0, fltIn_3 \in N \right\}$	0	N/A
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 \in D, fltIn_3 \in N \right\}$	$(-Inf, Inf)$	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 = 0, fltIn_2 \in D, fltIn_3 \in N \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in D, fltIn_2 \in D, fltIn_3 \in N \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 \in D, fltIn_3 \in N \right\}$	$(-Inf, Inf)$	N/A
fltIn1 infinity, *fltIn2* normalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in N, fltIn_3 \in N \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* normalized, i.e.		N/A

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in N, fltIn_3 \in N \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N, fltIn_3 \in N \end{array} \right\}$	$\langle -(2^{24} - 1), 2^{24} - 1 \rangle$	N/A
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, normalized or zero result and intermediate result i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M, \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right), \\ mr \subset N \cup \{-0, +0\} \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, result underflow i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0, \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, result overflow i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \in M \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{sign(refResult)•Inf, sign(refResult)•*pmax*}
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, intermediate result underflow i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ mr \in D \end{array} \right\}$	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-135. MLIB_Mac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, intermediate result overflow i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ mr \in M \end{array} \right\}$	$(-Inf, Inf)$	{Inf, -Inf}

3.50 Function MLIB_Mnac_FLT

3.50.1 Declaration

```
INLINEtFloat MLIB_Mnac_FLT(register tFloat fltIn1, register tFloat fltIn2, register tFloat
fltIn3);
```

3.50.2 Arguments

Table 3-136. MLIB_Mnac_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Input value to be subtracted.
register tFloat	fltIn2	**input**	First value to be multiplied.
register tFloat	fltIn3	**input**	Second value to be multiplied.

3.50.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2, fltIn3) \in X$ be a vector of inputs to MLIB_Mnac_FLT,

$mr = fltIn_2 \cdot fltIn_3$,

refResult be the theoretical exact result,

then

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs infinity, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid X \subset \{Inf, -Inf\} \right\}$	N/A	{Inf, *pmax*, -Inf, *nmax* }
fltIn1 zero or denormalized, *fltIn2* and *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D \bigvee fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{Inf, *pmax*} if sign(*fltIn2*)=sign(*fltIn3*), {-Inf, *nmax*} otherwise
fltIn1 normalized, *fltIn2* and *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	Inf if sign(*fltIn2*)=sign(*fltIn3*), -Inf otherwise
fltIn1 NaN, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = NaN \right\}$	$(-Inf, Inf)$	NaN
fltIn2 NaN, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_2 = NaN \right\}$	$(-Inf, Inf)$	NaN
fltIn3 NaN, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_3 = NaN \right\}$	$(-Inf, Inf)$	NaN
fltIn2 zero, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_2 = 0, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0}
fltIn2 infinity, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 = 0 \end{array} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0}
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, fltIn_2 \in D, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* infinity, i.e.	N/A	{NaN, Inf, -Inf, 0, -0}

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, fltIn_2 \in D, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$		
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, Inf, -Inf}
fltIn1 infinity, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	{NaN, Inf, -Inf}
fltIn1 zero, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{Inf, *pmax*} if sign(*fltIn2*)=sign(*fltIn3*), {-Inf, *nmax*} otherwise
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	N/A	{Inf, *pmax*} if sign(*fltIn2*)=sign(*fltIn3*), {-Inf, *nmax*} otherwise
fltIn1 normalized, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{array} \right\}$	$(-Inf, Inf)$	Inf if sign(*fltIn2*)=sign(*fltIn3*), -Inf otherwise
fltIn1 infinity, *fltIn2* zero, *fltIn3* zero, i.e.	N/A	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$		
All inputs zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 denormalized, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 = 0, fltIn_3 = 0 \end{array} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$	N/A	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* zero, i.e.	0	N/A

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle\|\; \begin{array}{c} fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 = 0 \end{array} \right\}$		
fltIn1 infinity, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle\|\; \begin{array}{c} fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{array} \right\}$	0	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle\|\; \begin{array}{c} fltIn_1 = 0, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{array} \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle\|\; \begin{array}{c} fltIn_1 \in D, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle\|\; \begin{array}{c} fltIn_1 \in N, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{array} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle\|\; \begin{array}{c} fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{array} \right\}$	N/A	{NaN, Inf, -Inf, *pmax*, *nmax*}

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 zero, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 = 0, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in D \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 denormalized, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in D, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in D \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 normalized, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in D \right\}$	$(-Inf, Inf)$	{NaN, Inf, -Inf, 0, -0}
fltIn1 infinity, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in \{Inf, -Inf\}, fltIn_2 = 0, fltIn_3 \in D \right\}$	N/A	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 = 0, fltIn_2 = 0, fltIn_3 \in D \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in D, fltIn_2 = 0, fltIn_3 \in D \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 = 0, fltIn_3 \in D \right\}$	0	N/A

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = \in D, fltIn_3 \in D \end{array} \right\}$	N/A	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in D, fltIn_3 \in D \end{array} \right\}$	0	N/A
All inputs denormalized, i.e. $\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid X \subset D, \}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* and *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 \in D \end{array} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, fltIn_3 \in D \end{array} \right\}$	N/A	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in N, fltIn_3 \in D \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in N, fltIn_3 \in D \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 normalized, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 \in N, fltIn_3 \in D \end{array} \right\}$	$(-Inf, Inf)$	N/A
fltIn1 infinity, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \end{array} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* infinity, *fltIn3* normalized, i.e.	N/A	{Inf, -Inf, *pmax*, *nmax*}

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \end{array} \right\}$		
fltIn1 denormalized, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \end{array} \right\}$	N/A	{Inf, -Inf, *pmax*, *nmax*}
fltIn1 normalized, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \end{array} \right\}$	$(-Inf, Inf)$	{Inf, -Inf}
fltIn1 infinity, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0, fltIn_3 \in N \end{array} \right\}$	N/A	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, \\ fltIn_2 = 0, fltIn_3 \in N \end{array} \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, \\ fltIn_2 = 0, fltIn_3 \in N \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 = 0, fltIn_3 \in N \end{array} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 \in N \end{array} \right\}$	$(-Inf, Inf)$	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in D, fltIn_3 \in N \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* normalized, i.e.	$\langle -0.5, 0.5 \rangle$	{-0, +0}

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in D, fltIn_3 \in N \end{array} \right\}$		
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 \in N \end{array} \right\}$	$(-Inf, Inf)$	N/A
fltIn1 infinity, *fltIn2* normalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, fltIn_3 \in N \end{array} \right\}$	N/A	{-sign(*fltIn1*)•Inf, -sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in N, fltIn_3 \in N \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N, fltIn_3 \in N \end{array} \right\}$	$\langle -(2^{24} - 1), 2^{24} - 1 \rangle$	N/A
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, normalized or zero result and intermediate result i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M, \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right), \\ mr \subset N \cup \{-0, +0\} \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, result underflow i.e.	$\langle -0.5, 0.5 \rangle$	{-0, +0}

Table continues on the next page...

Table 3-137. MLIB_Mnac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0, \end{matrix} \right\}$		
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, result overflow i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \in M \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	{sign(refResult)•Inf, sign(refResult)•*pmax*}
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, intermediate result underflow i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ mr \in D \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, intermediate result overflow i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ mr \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf}

3.51 Function MLIB_Msu_FLT

3.51.1 Declaration

```
INLINEtFloat MLIB_Msu_FLT(register tFloat fltIn1, register tFloat fltIn2, register tFloat
fltIn3);
```

3.51.2 Arguments

Table 3-138. MLIB_Msu_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Input value from which to subtract.
register tFloat	fltIn2	**input**	First value to be multiplied.
register tFloat	fltIn3	**input**	Second value to be multiplied.

3.51.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2, fltIn3) \in X$ be a vector of inputs to MLIB_Msu_FLT,

$mr = fltIn_2 \cdot fltIn_3$,

refResult be the theoretical exact result,

then

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs infinity, i.e. $\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid X \subset \{Inf, -Inf\} \}$	N/A	{Inf, *pmax*, -Inf, *nmax*, NaN}
fltIn1 zero or denormalized, *fltIn2* and *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D \bigvee fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	{-Inf, *nmax*} if sign(*fltIn2*)=sign(*fltIn3*), {Inf, *pmax*} otherwise
fltIn1 normalized, *fltIn2* and *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	-Inf if sign(*fltIn2*)=sign(*fltIn3*), Inf otherwise
fltIn1 NaN, i.e. $\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = NaN \}$	$(-Inf, Inf)$	NaN
fltIn2 NaN, i.e. $\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_2 = NaN \}$	$(-Inf, Inf)$	NaN
fltIn3 NaN, i.e. $\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_3 = NaN \}$	$(-Inf, Inf)$	NaN

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn2 zero, *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_2 = 0, fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0}
fltIn2 infinity, *fltIn3* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_2 \in \{Inf, -Inf\}, fltIn_3 = 0 \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, 0, -0}
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D, fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, fltIn_2 \in D, \\ fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, fltIn_2 \in D, \\ fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, Inf, -Inf}
fltIn1 infinity, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{matrix} \right\}$	$(-Inf, Inf)$	{NaN, Inf, -Inf}
fltIn1 zero, *fltIn2* normalized, *fltIn3* infinity, i.e.	N/A	{-Inf, *nmax*} if sign(*fltIn2*)=sign(*fltIn3*), {Inf, *pmax*} otherwise

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{\begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{matrix}\right\}$		
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{matrix}\right\}$	N/A	{-Inf, *nmax*} if sign(*fltIn2*)=sign(*fltIn3*), {Inf, *pmax*} otherwise
fltIn1 normalized, *fltIn2* normalized, *fltIn3* infinity, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, fltIn_2 \in N, \\ fltIn_3 \in \{Inf, -Inf\} \end{matrix}\right\}$	$(-Inf, Inf)$	-Inf if sign(*fltIn2*)=sign(*fltIn3*), Inf otherwise
fltIn1 infinity, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0, fltIn_3 = 0 \end{matrix}\right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
All inputs zero, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, fltIn_2 = 0, fltIn_3 = 0 \end{matrix}\right\}$	0	{-0, +0}
fltIn1 denormalized, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 = 0, fltIn_3 = 0 \end{matrix}\right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* zero, i.e. $\left\{\begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 = 0, fltIn_3 = 0 \end{matrix}\right\}$	0	N/A

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in D, fltIn_3 = 0 \end{matrix} \middle\vert \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \\ fltIn_1 = 0, \\ fltIn_2 \in D, fltIn_3 = 0 \end{matrix} \middle\vert \right\}$	0	{-0, +0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \\ fltIn_1 \in D, \\ fltIn_2 \in D, fltIn_3 = 0 \end{matrix} \middle\vert \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \\ fltIn_1 \in N, \\ fltIn_2 \in D, fltIn_3 = 0 \end{matrix} \middle\vert \right\}$	0	N/A
fltIn1 infinity, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{matrix} \middle\vert \right\}$	0	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \\ fltIn_1 = 0, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{matrix} \middle\vert \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* zero, i.e.	0	{-0, +0}

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; \begin{matrix} fltIn_1 \in D, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{matrix} \right\}$		
fltIn1 normalized, *fltIn2* normalized, *fltIn3* zero, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; \begin{matrix} fltIn_1 \in N, \\ fltIn_2 \in N, \\ fltIn_3 = 0 \end{matrix} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; \begin{matrix} fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; \begin{matrix} fltIn_1 = 0, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 denormalized, *fltIn2* infinity, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; \begin{matrix} fltIn_1 \in D, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{matrix} \right\}$	N/A	{NaN, Inf, -Inf, 0, -0}
fltIn1 normalized, *fltIn2* infinity, *fltIn3* denormalized, i.e.	$(-Inf, Inf)$	{NaN, Inf, -Inf, 0, -0}

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in N, \\ fltIn_2 \in \{Inf, -Inf\}, \\ fltIn_3 \in D \end{array} \right\}$		
fltIn1 infinity, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, \\ fltIn_2 = 0, fltIn_3 \in D \end{array} \right\}$	0	N/A
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in \{Inf, -Inf\}, \\ fltIn_2 = \in D, fltIn_3 \in D \end{array} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in D, fltIn_3 \in D \end{array} \right\}$	0	N/A
All inputs denormalized, i.e. $\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid X \subset D, \}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* and *fltIn3* denormalized, i.e.	0	N/A

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, fltIn_2 \in D, fltIn_3 \in D \right\}$		
fltIn1 infinity, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in N, fltIn_3 \in D \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, fltIn_2 \in N, fltIn_3 \in D \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, fltIn_2 \in N, fltIn_3 \in D \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 normalized, *fltIn2* normalized, *fltIn3* denormalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, fltIn_2 \in N, fltIn_3 \in D \right\}$	$(-Inf, Inf)$	N/A
fltIn1 infinity, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	N/A	{NaN, Inf, -Inf, 0, -0, *pmax*, *nmax*}
fltIn1 zero, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 = 0, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	N/A	{Inf, -Inf, *pmax*, *nmax*}
fltIn1 denormalized, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in D, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	N/A	{Inf, -Inf, *pmax*, *nmax*}
fltIn1 normalized, *fltIn2* infinity, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \in N, fltIn_2 \in \{Inf, -Inf\}, fltIn_3 \in N \right\}$	$(-Inf, Inf)$	{Inf, -Inf}

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 infinity, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in \{Inf, -Inf\}, fltIn_2 = 0, fltIn_3 \in N \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 = 0, fltIn_2 = 0, fltIn_3 \in N \right\}$	0	N/A
fltIn1 denormalized, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in D, fltIn_2 = 0, fltIn_3 \in N \right\}$	0	{-0, +0}
fltIn1 normalized, *fltIn2* zero, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 = 0, fltIn_3 \in N \right\}$	0	N/A
fltIn1 infinity, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 \in D, fltIn_3 \in N \right\}$	$(-Inf, Inf)$	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 = 0, fltIn_2 \in D, fltIn_3 \in N \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in D, fltIn_2 \in D, fltIn_3 \in N \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 normalized, *fltIn2* denormalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in N, fltIn_2 \in D, fltIn_3 \in N \right\}$	$(-Inf, Inf)$	N/A
fltIn1 infinity, *fltIn2* normalized, *fltIn3* normalized, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3 \in X \;\middle|\; fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in N, fltIn_3 \in N \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 zero, *fltIn2* normalized, *fltIn3* normalized, i.e.		N/A

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 = 0, \\ fltIn_2 \in N, fltIn_3 \in N \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	
fltIn1 denormalized, *fltIn2* normalized, *fltIn3* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid \\ fltIn_1 \in D, \\ fltIn_2 \in N, fltIn_3 \in N \end{array} \right\}$	$\langle -(2^{24} - 1), 2^{24} - 1 \rangle$	N/A
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, normalized or zero result and intermediate result i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M, \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right), \\ mr \subset N \cup \{-0, +0\} \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, result underflow i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0, \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, result overflow i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \in M \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{sign(refResult)•Inf, sign(refResult)•*pmax*}
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, intermediate result underflow i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ mr \in D \end{array} \right\}$	$(-Inf, Inf)$	N/A

Table continues on the next page...

Table 3-139. MLIB_Msu_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
fltIn1 finite, *fltIn2* and *fltIn3* normalized or denormalized, intermediate result overflow i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3 \in X \mid fltIn_1 \notin M \\ fltIn_2 \in N \cup D, fltIn_3 \in N \cup D, \\ mr \in M \end{matrix} \right\}$	$(-Inf, Inf)$	{Inf, -Inf}

3.52 Function MLIB_Mul_FLT

3.52.1 Declaration

```
INLINEtFloat MLIB_Mul_FLT(register tFloat fltIn1, register tFloat fltIn2);
```

3.52.2 Arguments

Table 3-140. MLIB_Mul_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Operand is a single precision floating point number.
register tFloat	fltIn2	**input**	Operand is a single precision floating point number.

3.52.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2) \in X$ be a vector of inputs to MLIB_Mul_FLT,

refResult be the theoretical exact result,

then

Table 3-141. MLIB_Mul_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Both inputs infinity, i.e.	N/A	{Inf, -Inf, *pmax*, *nmax*}

Table continues on the next page...

Table 3-141. MLIB_Mul_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in \{Inf, -Inf\} \right\}$		
fltIn1 normalized number, NaN, or infinity, *fltIn2* NaN, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in N \cup NaN \cup \{Inf, -Inf\}, fltIn_2 = NaN \right\}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 denormalized or zero, *fltIn2* NaN, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D \cup \{-0, +0\}, fltIn_2 = NaN \right\}$	N/A	{NaN, -0, +0}
fltIn1 infinity, *fltIn2* denormalized or zero, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\} fltIn_1 \in D \cup \{-0, +0\} \right\}$	N/A	{NaN, Inf, -Inf, -0, +0}
fltIn1 infinity, *fltIn2* normalized, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\} fltIn_1 \in N \right\}$	N/A	{Inf, -Inf, *pmax*, *nmax*}
fltIn1 NaN, *fltIn2* infinity or normalized, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN, fltIn_2 \in N \cup \{Inf, -Inf\} \right\}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 NaN, *fltIn2* denormalized or zero, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN, fltIn_2 \in D \cup \{-0, +0\} \right\}$	N/A	{NaN, -0, +0}
fltIn1 denormalized or zero, *fltIn2* infinity, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D \cup \{-0, +0\}, fltIn_2 \in \{Inf, -Inf\} \right\}$	N/A	{NaN, Inf, -Inf, -0, +0}
fltIn1 normalized, *fltIn2* infinity, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in N, fltIn_2 \in \{Inf, -Inf\} \right\}$	N/A	{Inf, -Inf, *pmax*, *nmax*}
Both inputs denormalized, i.e.	$\langle -0.5, 0.5 \rangle$	{-0, +0}

Table continues on the next page...

Table 3-141. MLIB_Mul_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\{ fltIn_1, fltIn_2 \in X \mid X \subset D \}$		
fltIn1 denormalized, *fltIn2* zero, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, fltIn_2 = 0 \}$	0	N/A
fltIn1 denormalized, *fltIn2* normalized, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in N \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 zero, *fltIn2* denormalized, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 \in D \}$	0	N/A
Both inputs zero, i.e. $\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 = 0 \}$	0	N/A
fltIn1 zero, *fltIn2* normalized, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in N \end{matrix} \right\}$	0	N/A
fltIn1 normalized, *fltIn2* denormalized, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N, fltIn_2 \in D \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 normalized, *fltIn2* zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N, fltIn_2 = 0 \end{matrix} \right\}$	0	N/A
Both inputs normalized, result normalized, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
Both inputs normalized, result overflow, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) = Inf \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	{Inf, -Inf, *pmax*, *nmax*}
Both inputs normalized, result underflow, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}

3.53 Function MLIB_Neg_FLT

3.53.1 Declaration

```
INLINEtFloat MLIB_Neg_FLT(register tFloat fltIn);
```

3.53.2 Arguments

Table 3-142. MLIB_Neg_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn	**input**	Input value which negative value should be returned.

3.53.3 Worst-Case Error Bounds

Let $(fltIn) \in X$ be an input to MLIB_Neg_FLT,

then

Table 3-143. MLIB_Neg_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Input infinity, i.e. $\{fltIn \in X \mid fltIn \in \{Inf, -Inf\}\}$	N/A	{-sign(*fltIn*)•Inf, -sign(*fltIn*)•*pmax*}
Input NaN, i.e. $\{fltIn \in X \mid fltIn = NaN\}$	N/A	{NaN, *pmax*, *nmax*}
Denormalized input, i.e. $\{fltIn \in X \mid X \subset D\}$	0	{-0, +0}
Normalized input, i.e. $\{fltIn \in X \mid X \subset N\}$	0	N/A
Zero input, i.e. $\{fltIn \in X \mid fltIn = 0\}$	0	N/A

3.54 Function MLIB_Sub_FLT

3.54.1 Declaration

```
INLINEtFloat MLIB_Sub_FLT(register tFloat fltIn1, register tFloat fltIn2);
```

3.54.2 Arguments

Table 3-144. MLIB_Sub_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	Operand is a single precision floating point number.
register tFloat	fltIn2	**input**	Operand is a single precision floating point number.

3.54.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2) \in X$ be a vector of inputs to MLIB_Sub_FLT,

then

Table 3-145. MLIB_Sub_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Both inputs infinity, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in \{Inf, -Inf\} \right\}$	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*}
fltIn1 arbitrary value, *fltIn2* NaN, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_2 = NaN \right\}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 infinity, *fltIn2* normalized, denormalized, or zero, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 \in \{Inf, -Inf\}, fltIn_2 \in N \cup D \cup \{-0, +0\} \right\}$	N/A	{sign(*fltIn1*)•Inf, sign(*fltIn1*)•*pmax*}
fltIn1 NaN, *fltIn2* arbitrary value, i.e. $\left\{ fltIn_1, fltIn_2 \in X \mid fltIn_1 = NaN \right\}$	N/A	{NaN, *pmax*, *nmax*}
fltIn1 normalized, denormalized, or zero, *fltIn2* infinity, i.e.	N/A	{-sign(*fltIn2*)•Inf, -sign(*fltIn2*)•*pmax*}

Table continues on the next page...

Table 3-145. MLIB_Sub_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N \cup D \cup \{-0, +0\}, \\ fltIn_2 \in \{Inf, -Inf\} \end{array} \right\}$		
Both inputs denormalized, i.e. $\{fltIn_1, fltIn_2 \in X \mid X \subset D\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* zero, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, fltIn_2 = 0\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
fltIn1 denormalized, *fltIn2* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid fltIn_1 \in D, \\ fltIn_2 \in N \end{array} \right\}$	$\langle -(2^{24} - 1), 2^{24} - 1 \rangle$	N/A
fltIn1 zero, *fltIn2* denormalized, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 \in D\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}
Both inputs zero, i.e. $\{fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, fltIn_2 = 0\}$	0	N/A
fltIn1 zero, *fltIn2* normalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid fltIn_1 = 0, \\ fltIn_2 \in N \end{array} \right\}$	0	N/A
fltIn1 normalized, *fltIn2* denormalized, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N, fltIn_2 \in D \end{array} \right\}$	$\langle -(2^{24} - 1), 2^{24} - 1 \rangle$	N/A
fltIn1 normalized, *fltIn2* zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid \\ fltIn_1 \in N, fltIn_2 = 0 \end{array} \right\}$	0	N/A
Both inputs normalized, result normalized or zero, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ refResult = 0 \vee \\ \left(\begin{array}{c} \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) \neq Inf \end{array} \right) \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	N/A
Both inputs normalized, result overflow, i.e.	$\langle -0.5, 0.5 \rangle$	{sign(refResult)•Inf, sign(refResult)•*pmax*}

Table continues on the next page...

Table 3-145. MLIB_Sub_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
$\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert + 0.5 \cdot \mathrm{ULP}(refResult) = Inf \end{array} \right\}$		
Both inputs normalized, result underflow, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2 \in X \mid X \subset N, \\ \lvert refResult \rvert - 0.5 \cdot \mathrm{ULP}(refResult) \in D \wedge \\ refResult \neq 0 \end{array} \right\}$	$\langle -0.5, 0.5 \rangle$	{-0, +0}

3.55 Function MLIB_VMac_FLT

3.55.1 Declaration

```
INLINEtFloat MLIB_VMac_FLT(register tFloat fltIn1, register tFloat fltIn2, register tFloat
fltIn3, register tFloat fltIn4);
```

3.55.2 Arguments

Table 3-146. MLIB_VMac_FLT arguments

Type	Name	Direction	Description
register tFloat	fltIn1	**input**	First input value to first multiplication.
register tFloat	fltIn2	**input**	Second input value to first multiplication.
register tFloat	fltIn3	**input**	First input value to second multiplication.
register tFloat	fltIn4	**input**	Second input value to second multiplication.

3.55.3 Worst-Case Error Bounds

Let $(fltIn1, fltIn2, fltIn3, fltIn4) \in X$ be a vector of inputs to MLIB_VMac_FLT,

$ir_1 = fltIn_1 \cdot fltIn_2,$

$ir_2 = fltIn_3 \cdot fltIn_4,$

refResult be the theoretical exact result,

then

Table 3-147. MLIB_VMac_FLT Worst-Case Error Bounds - Return Value

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
Any of the inputs is NaN or infinity, i.e. $\left\{ fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid X \cap M \neq \varnothing \right\}$	N/A	{Inf, -Inf, NaN, *pmax*, *nmax*, -0, +0}
Any of the inputs is denormalized, neither is NaN or infinity, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid \\ X \cap D \neq \varnothing, X \cap M = \varnothing \end{matrix} \right\}$	$(-Inf, Inf)$	N/A
All inputs normalized or zero, results normalized or zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ ir_1 \in N \cup \{-0, +0\}, ir_2 \in N \cup \{-0, +0\}, \\ refResult = 0 \vee \\ \left(\begin{matrix} \lvert refResult \rvert - \mathrm{ULP}(refResult) \notin D \wedge \\ \lvert refResult \rvert + \mathrm{ULP}(refResult) \neq Inf \end{matrix} \right) \end{matrix} \right\}$	$\langle -1, 1 \rangle$	N/A
All inputs zero, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid \\ X \subset \{-0, +0\} \end{matrix} \right\}$	0	N/A
All inputs normalized or zero, result overflow, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \lvert refResult \rvert + \mathrm{ULP}(refResult) = Inf \end{matrix} \right\}$	$\langle -1, 1 \rangle$	{sign(refResult)•Inf, sign(refResult)•*pmax*, NaN}
All inputs normalized or zero, result underflow or underflow in internal compensation calculation, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ 0.5 \cdot (\lvert refResult \rvert - \mathrm{ULP}(refResult) \in D) \wedge \\ refResult \neq 0 \end{matrix} \right\}$	$\langle -2^{24}, 2^{24} \rangle$	{-0, +0}
All inputs normalized or zero, intermediate result overflow, i.e. $\left\{ \begin{matrix} fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ \lvert ir_1 \rvert \in \{Inf, -Inf\} \vee \lvert ir_2 \rvert \in \{Inf, -Inf\} \end{matrix} \right\}$	N/A	Arbitrary floating-point value (including Inf, –Inf, NaN)

Table continues on the next page...

Table 3-147. MLIB_VMac_FLT Worst-Case Error Bounds - Return Value (continued)

Subset of input domain	Worst-case return value error bounds [ulp]	Allowed specific return values (regardless the error bounds)
All inputs normalized or zero, intermediate result underflow, i.e. $\left\{ \begin{array}{c} fltIn_1, fltIn_2, fltIn_3, fltIn_4 \in X \mid \\ X \subset N \cup \{-0, +0\}, \\ ir_1 \in D \vee ir_2 \in D \end{array} \right\}$	$(-Inf, Inf)$	N/A

How to Reach Us:

Home Page:
nxp.com

Web Support:
nxp.com/support

Document Number S32K14XMCFLTACC
Revision 3